

CONCENTRATION COMPACTNESS FOR CRITICAL RADIAL WAVE MAPS

ELISABETTA CHIODAROLI, JOACHIM KRIEGER, AND JONAS LÜHRMANN

ABSTRACT. We consider radially symmetric, energy critical wave maps from $(1 + 2)$ -dimensional Minkowski space into the unit sphere \mathbb{S}^m , $m \geq 1$, and prove global regularity and scattering for classical smooth data of finite energy. In addition, we establish a priori bounds on a suitable scattering norm of the radial wave maps and exhibit concentration compactness properties of sequences of radial wave maps with uniformly bounded energies. This extends and complements the beautiful classical work of Christodoulou-Tahvildar-Zadeh [3, 4] and Struwe [31, 33] as well as of Nahas [22] on radial wave maps in the case of the unit sphere as the target. The proof is based upon the concentration compactness/rigidity method of Kenig-Merle [6, 7] and a “twisted” Bahouri-Gérard type profile decomposition [1], following the implementation of this strategy by the second author and Schlag [17] for energy critical wave maps into the hyperbolic plane as well as by the last two authors [16] for the energy critical Maxwell-Klein-Gordon equation.

CONTENTS

1. Introduction	1
2. Preliminaries	10
3. Function spaces	11
4. Multilinear estimates	17
5. Decomposition of the nonlinearity and renormalization	20
6. Breakdown criterion	29
7. Concentration compactness step	40
8. Rigidity argument	96
References	102

1. INTRODUCTION

We consider wave maps $\phi: \mathbb{R}^{1+2} \rightarrow \mathbb{S}^m$ from $(1 + 2)$ -dimensional Minkowski space \mathbb{R}^{1+2} into the m -dimensional unit sphere $\mathbb{S}^m \hookrightarrow \mathbb{R}^{m+1}$, $m \geq 1$, satisfying the equation

$$(WM) \quad \square \phi = -\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi$$

for radially symmetric initial data

$$(\phi, \partial_t \phi)|_{t=0} = (\phi_0, \phi_1): \mathbb{R}^2 \rightarrow T\mathbb{S}^m.$$

Here we use the extrinsic formulation of the wave maps equation, viewing the \mathbb{R}^{m+1} -valued function ϕ as a column vector and denoting its transpose by ϕ^\dagger . In particular, we note that any initial data (ϕ_0, ϕ_1) must satisfy the consistency conditions

$$\phi_0^\dagger \phi_0 = 1, \quad \phi_0^\dagger \phi_1 = 0.$$

Greek indices such as α are implicitly assumed to run from 0 to 2 and we use the standard conventions for summing and raising/lowering indices with respect to the metric $\text{diag}(-1, +1, +1)$ on Minkowski space \mathbb{R}^{1+2} . We denote the d’Alembertian by $\square = -\partial_t^2 + \Delta$ and introduce the shorthand notation $\phi[t] = (\phi(t), \partial_t \phi(t))$ for $t \in \mathbb{R}$. An initial data pair $(\phi_0, \phi_1): \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ is called classical if it is smooth and constant in the exterior of a compact set.

The wave maps equation (WM) admits a non-negative conserved energy functional

$$E[\phi] := \frac{1}{2} \int_{\mathbb{R}^2} \sum_{\alpha=0}^2 |\partial_\alpha \phi|^2 dx,$$

which is invariant under the scaling of the equation

$$\phi(t, x) \mapsto \phi(\lambda t, \lambda x), \quad \lambda > 0.$$

The Cauchy problem for (WM) on \mathbb{R}^{1+2} is therefore energy critical.

The main result of this article asserts that for radially symmetric, classical initial data, the unique, smooth solutions to (WM) exist globally in time and scatter to finite energy free waves. Moreover, we establish a priori bounds on a suitable scattering norm of the solutions and we exhibit concentration compactness properties of sequences of radially symmetric wave maps into the unit sphere with uniformly bounded energies.

Theorem 1.1. *There exists a non-decreasing function $K: [0, \infty) \rightarrow [0, \infty)$ with the following property: Let (ϕ_0, ϕ_1) be radially symmetric, classical initial data of energy E . Then there exists a global, unique, smooth solution $\phi: \mathbb{R}^{1+2} \rightarrow \mathbb{S}^m$ to (WM) with initial data $\phi[0] = (\phi_0, \phi_1)$ satisfying the a priori bound*

$$\|\phi\|_S \leq K(E),$$

where the S norm is defined in Section 3 below. In particular, ϕ scatters to finite energy free waves as $t \rightarrow \pm\infty$ in the sense that there exist $(f_\pm, g_\pm) \in \dot{H}_x^1 \times L_x^2$ such that

$$\lim_{t \rightarrow \pm\infty} \|\nabla_{t,x} \phi - \nabla_{t,x} S(t)(f_\pm, g_\pm)\|_{L_x^2} = 0,$$

where $S(t)(f_\pm, g_\pm) = \cos(t|\nabla|)f_\pm + \frac{\sin(t|\nabla|)}{|\nabla|}g_\pm$ denotes the free wave propagator.

We emphasize that global regularity and scattering for energy critical radial wave maps into arbitrary compact target manifolds has already been established in by now classical works of Christodoulou-Tahvildar-Zadeh [3, 4] and Struwe [31, 33] as well as of Nahas [22]. We shall next give a brief overview of the history of the wave maps problem and motivate why we are revisiting the beautiful classical results on radial wave maps. Then we conclude this introduction with an overview of the proof of Theorem 1.1.

1.1. History and motivation. We note that the wave maps problem has been the subject of a fascinating and vast body of literature over the past decades that we cannot adequately review here in its entirety. Our primary focus shall be on energy critical wave maps from $(1+2)$ -dimensional Minkowski space.

The study of energy critical radial wave maps was begun in the seminal work of Christodoulou-Tahvildar-Zadeh [4] where global regularity is proven for arbitrary compact target manifolds for radially symmetric initial data with sufficiently small energy. This small energy global regularity result is then strengthened to a large energy global regularity result by excluding concentration of energy provided the target manifold, unlike the sphere, satisfies a suitable convexity condition. In a subsequent paper [3], Christodoulou-Tahvildar-Zadeh also obtain pointwise scattering bounds on the solutions under the same convexity condition on the target manifold. Struwe [31, 33] then established large energy global regularity for radially symmetric wave maps into arbitrary compact target manifolds by combining the results from [4] with a careful blowup analysis from [32]. More recently, Nahas [22] also proved scattering for radially symmetric wave maps into arbitrary compact target manifolds.

For wave maps without any symmetry assumptions the sub-critical local well-posedness theory was developed by Klainerman-Machedon [8–10] and Klainerman-Selberg [12, 13], making crucial use of the null structure in the quadratic derivative nonlinearity of the wave maps equation.

A very important step in extending the small energy global regularity result of Christodoulou-Tahvildar-Zadeh [4] to the non-radial case was achieved by Tataru [38], establishing global regularity for energy critical

wave maps into arbitrary compact target manifolds for initial data which is small in the scale-invariant homogeneous Besov space $\dot{B}_{2,1}^1(\mathbb{R}^2) \times \dot{B}_{2,1}^0(\mathbb{R}^2)$. This work introduced an important functional framework for the study of the energy critical wave maps equation, in particular the delicate null frame spaces. The final breakthrough to prove small energy global regularity for energy critical wave maps into the unit sphere \mathbb{S}^m , $m \geq 1$, was achieved by Tao [37] through the key realization that certain non-perturbative terms in the wave maps nonlinearity can be cast into a better form by exploiting the gauge invariance of the wave maps problem. Small energy global regularity was then extended to other target manifolds by the second author [15] for the hyperbolic plane \mathbb{H}^2 and by Tataru [39] for arbitrary target manifolds that can be isometrically embedded into Euclidean space. The key role that the gauge structure plays in the study of the wave maps equation at the critical regularity to renormalize the equation into a better form was also further clarified in the works of Klainerman-Rodnianski [11], Shatah-Struwe [27], Nahmod-Stefanov-Uhlenbeck [23] and the second author [14] on global regularity for wave maps from higher-dimensional Minkowski space for small critical Sobolev data.

For large energies, depending on the geometry of the target manifold, blowup can occur for energy critical wave maps. Indeed, the blowup analysis of Struwe [32] for energy critical equivariant wave maps showed that singularity formation must be tied to the existence of non-trivial finite energy harmonic maps from \mathbb{R}^2 into the target manifold. Later, the second author joint with Schlag and Tataru [18], Raphaël-Rodnianski [24] and Rodnianski-Sterbenz [25] constructed examples of equivariant wave maps into \mathbb{S}^2 that blow up in finite time via the concentration of a non-trivial harmonic map.

These developments culminated in the threshold conjecture that for energy critical wave maps global regularity is expected to hold for initial data with energy less than the energy of any non-trivial harmonic map into the target manifold. This conjecture was established independently around the same time by the second author and Schlag [17] for the hyperbolic plane \mathbb{H}^2 as the target, by Tao [34] for all hyperbolic spaces \mathbb{H}^d , $d \geq 1$, and by Sterbenz-Tataru [29, 30] for any target manifold that can be isometrically embedded into Euclidean space.

Our motivation for this article is essentially twofold. On the one hand we analyze to what extent the complicated function spaces introduced in the seminal works of Tataru [38] and Tao [37] can be replaced by a simpler functional framework in the radial context to settle the perturbative theory for the wave maps equation (WM) similarly to the strategy in Tao [37]. Here our goal was to avoid Fourier localization as much as possible and apart from a basic spatial frequency localization, our framework in particular avoids modulation localizations as in [38] and [37]. More specifically, our space S to control the regularity of radial wave maps is built from dyadic subspaces S_k in the sense that

$$\|\phi\|_S^2 := \sum_{k \in \mathbb{Z}} \|\phi_k\|_{S_k}^2.$$

The dyadic subspace S_k is defined entirely in physical space and essentially has three parts of the following schematic form

$$\|\phi\|_{S_k} = \|\phi_k\|_{Str} + \sup_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_k\|_{L_t^2 L_x^2} + \sum_{\pm} \|(\partial_t \pm \partial_r) \phi_k\|_{Z_k^\pm},$$

see Definition 3.1 for the precise definition. The first part of the S_k norm consists of Strichartz-type norms where we crucially exploit the larger range of admissible Strichartz pairs in the radial context, see for instance Sterbenz [28] and Fang-Wang [5]. The second part is a local energy decay norm which quite naturally replaces the important $\dot{X}^{1, \frac{1}{2}, \infty}$ -type space from the non-radial context. Finally, the third part of the S_k norm is formed by certain atomic spaces Z_k^\pm . These basically correspond to an “incoming-outgoing” decomposition of free radial waves and involve $L_{t \pm r}^2 L_{t \mp r}^\infty$ -type norms which naturally replace the more complicated null frame spaces [38] from the non-radial context.

On the other hand, in this work we go beyond the results of Christodoulou-Tahvildar-Zadeh [3, 4], Struwe [31, 33] and Nahas [22] in the case of radially symmetric wave maps into the unit sphere \mathbb{S}^m ,

$m \geq 1$, as the target manifold and prove, in addition to global regularity and scattering, a priori bounds on the scattering norm S of the solutions to (WM) and we exhibit concentration compactness properties of sequences of radially symmetric wave maps into the unit sphere with uniformly bounded finite energies. To this end we use a version of the concentration compactness/rigidity method of Kenig-Merle [6, 7] and a modified Bahouri-Gérard type profile decomposition [1], following the implementation of this strategy by the second author and Schlag [17] for (non-radial) energy critical wave maps into the hyperbolic plane \mathbb{H}^2 and the implementation by the last two authors [16] for the related energy critical Maxwell-Klein-Gordon equation. Executing the concentration compactness/rigidity strategy in the context of energy critical wave maps is compounded by the presence of non-perturbative terms in the wave maps nonlinearity, which have to be dealt with via renormalization, and by certain strong low-high interactions in the wave maps nonlinearity. These difficulties will be explained in more detail in the next subsection. However, at this point we emphasize a key difference between our work for the unit sphere $\mathbb{S}^m \hookrightarrow \mathbb{R}^{m+1}$, $m \geq 1$, as the target and the work of the second author and Schlag [17] for the \mathbb{H}^2 target regarding the renormalization procedure. The fact that the gauge group for the target \mathbb{H}^2 is abelian was exploited heavily in [17] in order to implement a global-in-frequency gauge change by passing to the Coulomb gauge in the intrinsic setting. The construction of the Coulomb gauge is elementary and explicit in the abelian case, but it becomes problematic for large energies in the non-abelian case. Since in our work the gauge group is no longer abelian for the targets \mathbb{S}^m with $m \geq 3$, we instead employ a version of the gauge construction of Sterbenz-Tataru [29] in the extrinsic setting which deals separately with each frequency level. The latter is a refinement and further development of the microlocal gauge introduced by Tao [36, 37] and is better suited for large energies.

1.2. Overview of the proof of Theorem 1.1. Here we provide an outline of the main ideas of the proof of Theorem 1.1. Our goal is to show that there exists a non-decreasing function $K: [0, \infty) \rightarrow [0, \infty)$ with the following property: Let $(\phi_0, \phi_1): \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ be radially symmetric, classical initial data of energy E . Then there exists a global, unique, smooth wave map $\phi: \mathbb{R}^{1+2} \rightarrow \mathbb{S}^m$ with initial data $\phi[0] = (\phi_0, \phi_1)$ satisfying the a priori bound

$$\|\phi\|_S \leq K(E).$$

Once we have established this a priori bound, the scattering assertion of Theorem 1.1 is an immediate consequence.

Following the general philosophy of the concentration compactness/rigidity scheme, we argue by contradiction and assume that Theorem 1.1 fails. Then the existence of the function $K(\cdot)$ yielding the a priori bounds must fail at some finite energy level. Correspondingly, the following set of energies must be non-empty

$$\mathcal{E} := \left\{ E : \sup_{\{\phi: E[\phi] \leq E\}} \|\phi\|_{S[I]} = +\infty \right\},$$

where the supremum is taken over all radially symmetric wave maps $\phi: I \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$ with classical radially symmetric initial data of energy $E[\phi] \leq E$ and defined on some time interval I . We shall prove a small energy global regularity result together with a priori bounds on the scattering norm S for radial solutions to (WM) in Theorem 5.4. For this reason the infimum of the set \mathcal{E} has to be strictly positive

$$E_{crit} := \inf \mathcal{E} > 0.$$

Thus, we may pick a sequence of radial smooth wave maps $\phi^n: I^n \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$, $n \geq 1$, with maximal intervals of existence I^n such that

$$\lim_{n \rightarrow \infty} E[\phi^n] = E_{crit}, \quad \lim_{n \rightarrow \infty} \|\phi^n\|_{S[I^n]} = +\infty.$$

In the following we call such a sequence of wave maps essentially singular. Our goal is now to rule out the existence of such an essentially singular sequence of wave maps, hence proving Theorem 1.1. This will be achieved in the following two main steps.

- Extracting an energy class, radially symmetric, minimal blowup solution ϕ^∞ to (WM) of energy E_{crit} from the essentially singular sequence $\{\phi^n\}_{n \geq 1}$. As a key tool we use a “twisted” profile decomposition that takes into account the strong low-high interactions in the wave maps nonlinearity. We develop this modified Bahouri-Gérard type nonlinear profile decomposition for our setting by following the procedure introduced by the second author and Schlag [17] for energy critical wave maps into \mathbb{H}^2 . The minimal blowup solution can be thought of as a “minimal counterexample” to Theorem 1.1 and its orbit must therefore possess a strong compactness property modulo the symmetries of the equation. We note that the heart of this paper resides in this extraction procedure which is carried out in Section 7. All sections leading up to it lay the groundwork.
- Ruling out the minimal blowup solution ϕ^∞ via a version of the Kenig-Merle rigidity argument [7] as in [17]. Here we utilize the strong compactness property of ϕ^∞ against general properties of radial wave maps into the unit sphere. In particular, we invoke the non-existence of non-trivial, finite energy, radial harmonic maps into the unit sphere. This step is accomplished in Section 8.

The severe difficulties with the extraction of a minimal blowup solution to (WM) can be highlighted by comparing with the situation for the energy critical, defocusing nonlinear wave equation $\square u = u^5$ on \mathbb{R}^{1+3} . In this context Bahouri-Gérard [1] introduced a highly influential nonlinear profile decomposition. It basically asserts that a sequence of solutions to the quintic nonlinear wave equation with uniformly bounded energies can be decomposed into a sum of nonlinear solutions, which are referred to as the nonlinear profiles and which are rescaled and translated in space-time according to the non-compact symmetries of the equation, and an error term, which can be made small in a suitable norm. The ability to extract a minimal blowup solution ultimately relies on the asymptotic decoupling of different nonlinear profiles. In the quintic nonlinearity, the interactions of two different nonlinear profiles with essential frequency supports at divergent scales are asymptotically negligible. This reduces to consider diagonal frequency interactions. But then two different nonlinear profiles living at the same frequency scale must concentrate in divergent regions of space-time so that their interactions in the quintic nonlinearity again vanish asymptotically.

In contrast, for energy critical wave maps frequency diagonalization appears to partially fail at the critical regularity due to strong low-high interactions in the wave maps nonlinearity. In order to gain a better understanding of these difficulties in our context of radial wave maps into the unit sphere, we now take a closer look at the perturbative theory for the frequency localized wave maps equation (WM).

Renormalization and perturbative theory. More precisely, we study the evolution of a single dyadic frequency block ϕ_k , say $k = 0$, satisfying the equation

$$\square \phi_0 = -P_0(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi).$$

Upon decomposing each input of the nonlinearity into its Littlewood-Paley pieces, the nonlinear wave equation for ϕ_0 assumes the following form

$$\begin{aligned} \square \phi_0 = & -2\phi_{\leq -10} \partial_\alpha \phi_{\leq -10}^\dagger \partial^\alpha \phi_0 \\ & - \sum_{k_2 > -10} \sum_{k_3 = k_2 + O(1)} P_0(\phi_{\leq 20} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_{k_3}) \\ & - 2 \sum_{k_2 \leq -10} \phi_{\leq k_2 - 10} (P_0(\partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_{-5 < \cdot < 5}) - \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0) \\ & + \text{error}, \end{aligned}$$

where we are only listing those interaction terms that are more difficult to deal with and the other relatively insignificant interaction terms are just denoted by error. The difficulties with the high-high interactions of the second term and the third term on the right-hand side are a more peculiar feature of our framework. They are non-perturbative in the sense that they cannot be treated with the multilinear estimates within our functional framework due to its relatively elementary but appealing construction - for instance, we cannot

gain in the high-high interactions in our null form estimates. But we note that these two interaction terms could be easily handled with the full power of the more sophisticated functional framework from Tao [37]. Fortunately, this feature of our setting can be dealt with quite efficiently by passing to the “nonlinearly modified variable”

$$\bar{\phi}_0 := \phi_0 + \frac{1}{2} \sum_{k_2 > -10} \sum_{k_3 = k_2 + O(1)} P_0(\phi_{\leq 20} \phi_{k_2}^\dagger \phi_{k_3}) + \sum_{k_2 \leq -10} \phi_{\leq k_2 - 10} (P_0(\phi_{k_2}^\dagger \phi_{-5 < \cdot < 5}) - \phi_{k_2}^\dagger \phi_0),$$

which has the effect of either distributing derivatives to different inputs with a better frequency balance or of turning a trilinear interaction term into an easier quintilinear one upon reinserting the wave maps equation. The wave equation for the new variable $\bar{\phi}_0$ then takes on the form

$$\square \bar{\phi}_0 = -2\phi_{\leq -10} \partial_\alpha \phi_{\leq -10}^\dagger \partial^\alpha \phi_0 + \text{error}$$

and we are left to treat the more severe low-low-high interaction term on the right-hand side. Following Tao [36, 37] we exploit the geometry of the wave maps problem and use the orthogonality relation $\phi^\dagger \partial^\alpha \phi = 0$ which remains approximately preserved upon frequency localization. Then we arrive at a better equation for ϕ_0 , namely

$$(1.1) \quad \square \bar{\phi}_0 = -2(\phi_{\leq -10} \partial_\alpha \phi_{\leq -10}^\dagger - \partial_\alpha \phi_{\leq -10} \phi_{\leq -10}^\dagger) \partial^\alpha \phi_0 + \text{error},$$

where the matrix $(\phi_{\leq -10} \partial_\alpha \phi_{\leq -10}^\dagger - \partial_\alpha \phi_{\leq -10} \phi_{\leq -10}^\dagger)$ is anti-symmetric. Now we have the following trilinear estimate from Proposition 4.3 at our disposal

$$(1.2) \quad \|\phi_{k_1} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0\|_{L_t^1 L_x^2} \lesssim 2^{-\delta(k_1 - k_2)} \|\phi_{k_1}\|_{S_{k_1}} \|\phi_{k_2}\|_{S_{k_2}} \|\phi_0\|_{S_0}, \quad k_2 + O(1) \leq k_1 \leq 0,$$

for some absolute constant $\delta > 0$. The key exponential gain in this estimate allows us to handle those parts of the low-low-high interaction term where a derivative falls on the lowest frequency. At the same time we make the fundamental observation that we do not gain exponentially in the largest frequency difference in this estimate, resulting in the strong low-high interactions alluded to before. We note that the direct analogue of this estimate for the non-radial setting is due to Tao [37] and is much more difficult to achieve than within our functional framework for the radial case. In view of (1.2), we split the interaction term on the right-hand side of (1.1) into two parts

$$(1.3) \quad \square \bar{\phi}_0 = 2A_{\alpha; \leq -10} \partial^\alpha \phi_0 - 2 \sum_{k_2 \leq -10} \sum_{k_2 - 10 < k_1 \leq -10} (\phi_{k_1} \partial_\alpha \phi_{k_2}^\dagger - \partial_\alpha \phi_{k_2} \phi_{k_1}^\dagger) \partial^\alpha \phi_0 + \text{error},$$

where we are introducing the connection form

$$A_{\alpha; \leq -10} = - \sum_{k_2 \leq -10} (\phi_{\leq k_2 - 10} \partial_\alpha \phi_{k_2}^\dagger - \partial_\alpha \phi_{k_2} \phi_{\leq k_2 - 10}^\dagger).$$

While in the second interaction term on the right-hand side of (1.3) a derivative falls on the lowest frequency and this part can therefore be handled with the trilinear estimate, this is not the case for the first interaction term. Thus, the latter term is non-perturbative and has to be renormalized into a better form. Following the method first introduced by Tao [36, 37], we define suitable $SO(m+1)$ -valued gauge transformations $U_{\leq -10}$ and pass to the new variable $U_{\leq -10} \bar{\phi}_0$ satisfying the nonlinear wave equation

$$(1.4) \quad \begin{aligned} \square(U_{\leq -10} \bar{\phi}_0) &= U_{\leq -10} (\square \bar{\phi}_0 - 2A_{\alpha; \leq -10} \partial^\alpha \phi_0) \\ &\quad + 2(\partial_\alpha U_{\leq -10} + U_{\leq -10} A_{\alpha; \leq -10}) \partial^\alpha \phi_0 \\ &\quad + 2\partial_\alpha U_{\leq -10} \partial^\alpha (\bar{\phi}_0 - \phi_0) \\ &\quad + (\square U_{\leq -10}) \bar{\phi}_0, \end{aligned}$$

where the key difficulty now resides in the second interaction term on the right-hand side. More precisely, we implement a version of the gauge construction due to Sterbenz-Tataru [29], which is better suited for

large energies, and define the gauge transformations $U_{\leq h}$ for $h \in \mathbb{R}$ as solutions to the ODE

$$\frac{d}{dh} U_{\leq h} = U_{\leq h} B_h, \quad \lim_{h \rightarrow -\infty} U_{\leq h} = Id,$$

where the anti-symmetric matrix B_h is of the schematic form

$$B_h = \phi_{\leq h-10} \phi_h^\dagger - \phi_h \phi_{\leq h-10}^\dagger.$$

In particular, the anti-symmetry of B_h ensures that the gauge transformations $U_{\leq h}$ as solutions to the above ODE are exactly orthogonal. This choice of gauge transformation effectively transfers a derivative to the lowest frequency in the key difficult term $2(\partial_\alpha U_{\leq -10} + U_{\leq -10} A_{\alpha; \leq -10}) \partial^\alpha \phi_0$ so that it can basically be handled with the trilinear estimate (1.2). Finally, we will have to transfer the bounds on the variable $U_{\leq -10} \bar{\phi}_0$ back to the frequency localized wave map ϕ_0 . In particular, we note that we have to introduce certain frequency and spatial truncations in the definitions of the gauge transformations and the new variable $\bar{\phi}_0$ in order to have additional sources of smallness at our disposal in the large energy setting. These modifications will be explained as we go in the later sections. This essentially settles the perturbative theory for (WM).

We now describe the extraction procedure of the minimal blowup solution ϕ^∞ to (WM) which necessitates the development of a “twisted” Bahouri-Gérard type profile decomposition to take into account the effect of the strong low-high interactions described above. This undertaking is additionally compounded by the fact that we have to work at the level of the gauged variables $U_{\leq -10} \bar{\phi}_0$ because only these satisfy a nonlinear wave equation with good perturbative properties. As in [17] we use a finite induction on frequency procedure to carefully disentangle the low-high frequency interactions.

Decomposition into frequency atoms and evolving the lowest frequency non-atomic part. The first step consists in decomposing the essentially singular sequence of data $\{\phi^n[0]\}_{n \geq 1}$ into frequency atoms using the Métivier-Schochet procedure [21] as in Bahouri-Gérard [1]. Roughly speaking, the basic idea then goes as follows. Ultimately, we would like to conclude that upon passing to a subsequence, if necessary, the essentially singular sequence of data $\{\phi^n[0]\}_{n \geq 1}$ consists of exactly one frequency atom which in turn consists of exactly one concentration profile (to be defined precisely in Subsection 7.4) of asymptotic energy E_{crit} . In this scenario, the sequence $\{\phi^n[0]\}_{n \geq 1}$ has sufficient compactness properties that allow us to pass to a certain limit whose wave maps evolution will be the desired minimal blowup solution to (WM). In order to rule out all other possible scenarios, we seek to prove uniform in n , finite, global S norm bounds on (a subsequence of) the sequence of wave maps evolutions $\{\phi^n\}_{n \geq 1}$, which would contradict that the sequence is essentially singular. To this end we first achieve control over the wave maps evolutions of certain low frequency truncations of the essentially singular sequence of data $\{\phi^n[0]\}_{n \geq 1}$. Using a finite inductive procedure over the increasing size of the frequency supports of these low frequency truncations, we then conclude uniform in n , finite, global S norm bounds on the actual essentially singular sequence $\{\phi^n\}_{n \geq 1}$.

Using a version of the Métivier-Schochet procedure [21] we obtain a decomposition into frequency atoms (of a subsequence) of the essentially singular sequence of data

$$\phi^n[0] = \sum_{a=1}^{\Lambda} \phi^{na}[0] + \phi^{n\Lambda}[0],$$

where the frequency atoms $\phi^{na}[0]$ and the remainder term $\phi^{n\Lambda}[0]$ have disjoint frequency supports. The frequency atoms are sharply localized around frequency scales $(\lambda_n^a)^{-1}$, more precisely they have frequency support on $\{|\xi| \in [(\lambda_n^a)^{-1} R_n^{-1}, (\lambda_n^a)^{-1} R_n]\}$ for some sequence $R_n \rightarrow \infty$ growing sufficiently slowly. The frequency scales diverge from each other as $n \rightarrow \infty$ in the sense that

$$\lim_{n \rightarrow \infty} \frac{\lambda_n^a}{\lambda_n^b} + \frac{\lambda_n^b}{\lambda_n^a} = +\infty, \quad a \neq b.$$

We may assume that the atoms are ordered in terms of the increasing size of their frequency support scales $(\lambda_n^a)^{-1}$ and we introduce the notation $\mu_n^a = -\log(\lambda_n^a)$ for the corresponding dyadic frequency support scales. Moreover, the remainder term $\phi^{n\Lambda}[0]$ satisfies the important Besov norm smallness

$$\lim_{\Lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} \|\phi^{n\Lambda}[0]\|_{\dot{B}_{2,\infty}^1 \times \dot{B}_{2,\infty}^0} = 0.$$

As described above, ultimately we would like to conclude that there is exactly one atom in the decomposition (7.1), i.e. $\Lambda = 1$, which is of asymptotic energy E_{crit} . If this is the case, we proceed directly to the next stage below where we consider the evolution of the first “large” frequency atom. Otherwise, we now start a finite inductive procedure to conclude that the sequence $\{\phi^n\}_{n \geq 1}$ cannot be essentially singular. To this end we fix an integer Λ_0 sufficiently large such that upon passing to a subsequence, if necessary,

$$\sum_{a > \Lambda_0} \limsup_{n \rightarrow \infty} \|\phi^{na}[0]\|_{\dot{H}_x^1 \times L_x^2}^2 \leq \varepsilon_0,$$

where $\varepsilon_0 > 0$ is a sufficiently small constant that plays the role of a perturbative threshold in the key bootstrap argument in Proposition 7.9. In particular, ε_0 will be chosen sufficiently small depending only on the size of E_{crit} and it will be chosen to be less than the small energy global regularity threshold established in Theorem 5.4. Then we observe that due to the sharp frequency localizations of the atoms $\phi^{na}[0]$, $1 \leq a \leq \Lambda_0$, the remainder term $\phi^{n\Lambda_0}[0]$ gets split into $\Lambda_0 + 1$ “frequency shells”

$$\phi^{n\Lambda_0}[0] = \phi^{n\Lambda_0^{(0)}}[0] + \phi^{n\Lambda_0^{(1)}}[0] + \dots + \phi^{n\Lambda_0^{(\Lambda_0)}}[0],$$

where $\phi^{n\Lambda_0^{(0)}}[0]$ shall denote the lowest frequency component.

Our first step now consists in showing that the lowest frequency “non-atomic” component $\phi^{n\Lambda_0^{(0)}}[0]$ can be globally evolved and satisfies finite S norm bounds just in terms of E_{crit} uniformly for all sufficiently large n . Since the component $\phi^{n\Lambda_0^{(0)}}[0]$ may still have large energy, in order to be able to infer these S norm bounds, we approximate $\phi^{n\Lambda_0^{(0)}}[0]$ by a finite number of delicately chosen low frequency truncations $P_{\leq b_L} \phi^{n\Lambda_0^{(0)}}[0]$. However, up to this point we have totally ignored that the frequency truncations $P_{\leq b_L} \phi^{n\Lambda_0^{(0)}}[0] = P_{\leq b_L} \phi^n[0]$ are not “geometric” in the sense that they are not actual maps $\mathbb{R}^2 \rightarrow T\mathbb{S}^m$ into the unit sphere and therefore do not constitute suitable initial data for the wave maps equation (WM). To overcome this issue we just project the frequency truncations back to the sphere, using the normal projection operator, and denote the resulting initial data by $\Pi_{\leq b_L} \phi^n[0]$. This operation is well-defined and the frequency localization properties are approximately preserved up to exponential tails if around the frequency cut-offs a certain Besov norm smallness condition is satisfied, see Proposition 7.2. This, in particular, forces us to carry out a further refined frequency atom decomposition of the lowest frequency non-atomic part $\phi^{n\Lambda_0^{(0)}}[0]$ to carefully pick these low frequency cut-offs b_L . Using a finite induction procedure we then obtain uniform in n , a priori bounds on the S norms of the global evolutions of the “geometric” lowest frequency non-atomic parts $\Pi_{\leq \mu_n^1 - \log R_n} \phi^n[0]$ via an iterative bootstrap argument, which is accomplished in Proposition 7.8 in Subsection 7.3.

Selecting concentration profiles and adding the first large frequency atom. Having established control over the global evolution of the lowest frequency non-atomic part $\Pi_{\leq \mu_n^1 - \log(R_n)} \phi^n[0]$ in the previous step, we now “add in” the first large frequency atom $\phi^{n1}[0] = P_{[\mu_n^1 - \log(R_n), \mu_n^1 + \log(R_n)]} \phi^n[0]$ in the sense that we now attempt to globally evolve the geometric initial data

$$\Pi_{\leq \mu_n^1 + \log(R_n)} \phi^n[0].$$

In this paragraph we shall slightly abuse notation and write for simplicity

$$\phi^n[0] \equiv \Pi_{\leq \mu_n^1 + \log(R_n)} \phi^n[0],$$

denoting the evolution of this data by ϕ^n . By rescaling we may assume that $\mu_n^1 \equiv 0$ and we shall later denote by $\mu_n = \mu_n^1 - \log(R_n)$ the frequency cut-off delimiting the essential frequency supports of the lowest frequency non-atomic part and of the first large frequency atom. Moreover, we use the notation

$$\phi^n[0] \equiv \Pi_{\leq \mu_n^1 - \log(R_n)} \phi^n[0] + (\Pi_{\leq \mu_n^1 + \log(R_n)} \phi^n - \Pi_{\leq \mu_n^1 - \log(R_n)} \phi^n)[0] \equiv u^n[0] + \epsilon^n[0],$$

where u^n denotes the global evolution of the lowest frequency non-atomic part established in the previous stage. Since only the gauged variables satisfy a nonlinear wave equation with good perturbative properties, we now enact a Bahouri-Gérard type profile decomposition at the level of the gauged variables $U_{<k}^{(\phi^n)} \overline{\phi_k^n}$, which then has to be translated back to the level of the coordinate functions ϕ^n . Here, $U^{(\phi^n)}$ denote the gauge transformations defined by the evolutions ϕ^n . In order to take into account the strong coupling between the very low frequencies coming from u^n and the high frequencies, we have to extract the concentration profiles for ϵ^n with respect to a suitably modified linear wave operator to match the proper asymptotic evolution of the variables $U_{<k}^{(\phi^n)} \overline{\phi_k^n}$ as $t \rightarrow \pm\infty$. In view of (1.3) and (1.4) we use the following very natural linear magnetic-type wave equation to select our concentration profiles

$$(1.5) \quad \boxed{\square \tilde{\epsilon} = 2U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha, \text{low}}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \tilde{\epsilon}}$$

where

$$\mathcal{A}_{\alpha, \text{low}}(u^n) = - \sum_{k_2 < \mu_n} \sum_{k_2 - 10 < k_1 < \mu_n} (u_{k_1}^n) (\partial_\alpha u_{k_2}^n)^\dagger - (\partial_\alpha u_{k_2}^n) (u_{k_1}^n)^\dagger + (U_{<\mu_n}^{(u^n)})^\dagger \partial_\alpha U_{<\mu_n}^{(u^n)} + A_{\alpha; <\mu_n}^{(u^n)}.$$

The gauge transformations $U_{<\mu_n}^{(u^n)}$ and the anti-symmetric matrix $\mathcal{A}_{\alpha, \text{low}}(u^n)$ are defined purely in terms of the low frequency wave maps u^n over which we already have global control at this stage. Moreover, the anti-symmetry of $\mathcal{A}_{\alpha, \text{low}}(u^n)$ and the much lower essential frequency support of u^n are key for proving asymptotic energy conservation for the flow associated with (1.5), see Lemma 7.15. We note that the linear magnetic-type wave equation (1.5) is the direct analogue of the linear magnetic wave equation in [17, Definition 9.18] for the extraction of the concentration profiles in the context of energy critical wave maps into the hyperbolic plane \mathbb{H}^2 .

A pleasant feature of the linear magnetic-type wave equation (1.5) is that it is in fact independent of the frequency level $k \in \mathbb{Z}$ one works at. Roughly speaking, to obtain the profile decomposition at the level of ϕ^n one then evolves the data $\tilde{\epsilon}_k[0] := U_{<k}^{(\phi^n)} \epsilon_k^n[0]$ with respect to the flow of (1.5) for each frequency $k \in \mathbb{Z}$, selecting appropriate concentration times and profiles, and passes back to the level of ϕ^n by multiplying with $(U_{<k}^{(\phi^n)})^\dagger$. Summing over all frequencies $k \in \mathbb{Z}$ then roughly furnishes the desired profiles, see equation (7.38) and the remarks following it for the precise definitions. Provided that all concentration profiles have energy strictly less than E_{crit} we may then carefully construct the global evolution of the data $\Pi_{\leq \mu_n^1 + \log(R_n)} \phi^n[0]$ and obtain uniform in n , a priori bounds on the S norms, where a lot of work is required to accomplish that the profile decomposition ansatz is “sufficiently geometric”. This whole step is carried out in Theorem 7.17 in Subsection 7.4.

Conclusion of the induction on frequency process. We now continue this induction on frequency process and by proceeding as in Subsection 7.3 obtain that the data $\Pi_{\leq \mu_n^2 - \log(R_n)} \phi^n[0]$ can be globally evolved with uniform S norm bounds. Then we “add in” the second frequency atom $P_{[\mu_n^2 - \log(R_n), \mu_n^2 + \log(R_n)]} \phi^n[0]$ in the sense that by proceeding analogously to Subsection 7.4 we may establish the global evolution of the corresponding geometric data $\Pi_{\leq \mu_n^2 + \log(R_n)} \phi^n[0]$ with uniform S norm bounds under the assumption that the associated profiles all have energy strictly less than E_{crit} .

All in all, we may continue this procedure Λ_0 many times and establish the global evolution with uniform in n , a priori S norm bounds of (a subsequence of) the essentially singular sequence of data $\{\phi^n[0]\}_{n \geq 1}$, which would however be a contradiction, unless (a subsequence of) the sequence $\{\phi^n[0]\}_{n \geq 1}$ is composed of exactly one frequency atom that consists of exactly one profile of asymptotic energy E_{crit} . Thus, we must be in the latter scenario and (a subsequence of) $\{\phi^n[0]\}_{n \geq 1}$ has sufficient compactness properties that allow us

to pass to a certain limit whose wave maps evolution will be the desired minimal blowup solution to (WM) as detailed in Subsection 7.5. Here we note that the minimal blowup solution will merely have energy class regularity but that a strong local well-posedness theory for (WM) is only available at sub-critical regularities. For this reason we actually have to introduce a concept of energy class radial wave maps, which we achieve in Subsection 7.2 by regularization and reduction to the small energy case via finite speed of propagation, analogously to the procedures in [17] and [16].

The outcome of this last stage is that there exists a non-trivial, energy class, radially symmetric, minimal blowup solution $\phi^\infty: I \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$ to (WM) of energy E_{crit} and with maximal interval of existence I . Moreover, there exists a continuous function $\lambda: I \rightarrow (0, \infty)$ so that the family of functions

$$\{(\phi^\infty(t, \lambda(t)^{-1} \cdot), \lambda(t)^{-1} \partial_t \phi^\infty(t, \lambda(t)^{-1} \cdot)): t \in I\}$$

is pre-compact in $\dot{H}_x^1 \times L_x^2$.

Rigidity argument. Finally, we have to rule out the existence of such a minimal blowup solution ϕ^∞ to (WM). To this end we closely mimic the Kenig-Merle rigidity argument [7] as implemented in [17]. In particular, we invoke the non-existence of non-trivial, finite energy, radially symmetric harmonic maps into the unit sphere. This finishes the outline of the main ideas entering the proof of Theorem 1.1.

We expect that the restriction to the unit sphere \mathbb{S}^m , $m \geq 1$, as the target manifold in our work is not necessary and that our method extends to arbitrary compact target manifolds since a similar gauge construction as for the unit sphere works upon establishing additional estimates on the second fundamental form as in Tataru [39] and Sterbenz-Tataru [29].

Moreover, we point out that neither the gauge construction nor the selection process for the concentration profiles in our work crucially hinge on the radial symmetry assumption. We therefore suspect that the method of this article combined with the sophisticated functional framework from Tao [37] ought to allow for a similar result in the non-radial case upon restricting to energy levels below the energy of any non-trivial harmonic map from \mathbb{R}^2 to \mathbb{S}^m .

Organization of the paper. In Section 2 we introduce some notation and several basic definitions. In Section 3 we present our precise functional framework. In Section 4 we collect the most important multilinear estimates to handle the wave maps nonlinearity in the radial case. In Section 5 we carefully analyze the structure of the frequency localized wave maps nonlinearity and introduce the renormalization procedure to deal with the non-perturbative terms. Moreover, we establish a small energy global regularity result for (WM). In Section 6 we show that the S norm provides sufficient control on radial wave maps with classical initial data to infer long time existence and scattering. In Section 7 we begin with the actual proof of Theorem 1.1 and accomplish the most difficult step of extracting a minimal blowup solution with the strong compactness property. In Section 8 we rule out the existence of the minimal blowup solution and thus finish the proof of Theorem 1.1.

2. PRELIMINARIES

Notation and conventions. We write $A \lesssim B$ to denote $A \leq CB$ for some absolute constant $C > 0$ that may depend on fixed parameters and we shall use the notation $A \ll B$ to indicate that the implicit constant in the estimate is small. Moreover, we borrow from Tao [37] a convenient notation to describe multilinear expressions of product type. For scalar functions $\phi^{(1)}(t, x), \dots, \phi^{(n)}(t, x)$ we denote by $L(\phi^{(1)}, \dots, \phi^{(n)})(t, x)$ any multilinear expression of the form

$$L(\phi^{(1)}, \dots, \phi^{(n)})(t, x) := \int K(y_1, \dots, y_n) \phi^{(1)}(t, x - y_1) \cdots \phi^{(n)}(t, x - y_n) dy_1 \dots dy_n,$$

where the kernel K is a measure with bounded mass. We extend this notation to the case when $\phi^{(1)}, \dots, \phi^{(n)}$ take values as $(m + 1)$ -dimensional vectors or as $(m + 1) \times (m + 1)$ matrices.

Littlewood-Paley projections. We denote by φ a non-negative smooth cut-off function satisfying $\varphi(y) = 1$ for $y \leq 1$ and $\varphi(y) = 0$ for $y > 2$. Then we set $\varphi_0(y) = \varphi(y) - \varphi(2y)$ and $\varphi_k(y) = \varphi_0(2^{-k}y)$ for $k \in \mathbb{Z}$. We define the dyadic Littlewood-Paley projection operators P_k for $k \in \mathbb{Z}$ by

$$\widehat{P_k f}(\xi) = \varphi_k(|\xi|) \hat{f}(\xi).$$

We often write $f_k = P_k f$. Occasionally, we also need to use continuous Littlewood-Paley projections P_h for $h \in \mathbb{R}$. We recall the following Leibniz rule for the Littlewood-Paley projections P_k , see [37, Lemma 2].

Lemma 2.1. *It holds that*

$$(2.1) \quad P_k(fg) = f P_k g + L(\nabla_x f, 2^{-k} g).$$

Spatial cut-offs. We will also make use of cut-off functions for the radial variable $r \equiv |x|$. For $\ell \in \mathbb{Z}$ we denote by $\chi_{\{r \sim 2^\ell\}}$ a smooth non-negative bump function supported in $\{r \sim 2^\ell\}$ such that we have a smooth finite partition of unity

$$\sum_{\ell \in \mathbb{Z}} \chi_{\{r \sim 2^\ell\}}(r) = 1 \quad \text{for } r > 0.$$

Moreover, we denote by $\chi_{\{r \lesssim 2^\ell\}}$ a smooth cut-off function to $\{r \lesssim 2^\ell\}$. Analogously, we define the cut-off functions $\chi_{\{r > 2^\ell\}}$ and $\chi_{\{2^{\ell_1} \leq r \leq 2^{\ell_2}\}}$.

Frequency envelopes. We shall use the tool of frequency envelopes from [36] to track the frequency distribution of certain norms. A sequence $\{c_k\}_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$ of positive real numbers is a frequency envelope if

$$2^{-\sigma|k-k'|} c_{k'} \lesssim c_k \lesssim 2^{+\sigma|k-k'|} c_{k'}$$

for all $k, k' \in \mathbb{Z}$, where $\sigma > 0$ is a small absolute constant. We say that an initial data set $\phi[0]$ lies underneath the envelope $\{c_k\}_{k \in \mathbb{Z}}$ if

$$\|P_k \phi[0]\|_{\dot{H}_x^1 \times L_x^2} \leq c_k$$

for all $k \in \mathbb{Z}$. Given an initial data set $\phi[0]$ we may obtain such an envelope by defining

$$c_k := \sum_{\ell \in \mathbb{Z}} 2^{-\sigma|k-\ell|} \|P_\ell \phi[0]\|_{\dot{H}_x^1 \times L_x^2}.$$

3. FUNCTION SPACES

In this section we introduce the functional framework used for the proof of Theorem 1.1. The space S containing the radial wave map ϕ is built from dyadic subspaces S_k in the sense that

$$\|\phi\|_S^2 := \sum_{k \in \mathbb{Z}} \|\phi_k\|_{S_k}^2,$$

where the space S_k is defined as follows.

Definition 3.1. *Let $k \in \mathbb{Z}$ and let ϕ be a radially symmetric function on \mathbb{R}^{1+2} with Fourier support in $\{|\xi| \sim 2^k\}$. Then we define*

$$\begin{aligned} \|\phi\|_{S_k} &:= \|\nabla_{t,x} \phi\|_{L_t^\infty L_x^2} + \sup_{2 \leq q, p \leq \infty, \frac{1}{q} + \frac{1}{p} < \frac{1}{2}} 2^{(\frac{1}{q} + \frac{2}{p} - 1)k} \|\nabla_{t,x} \phi\|_{L_t^q L_x^p} + \sup_{0 < \lambda < 1} 2^{(\frac{1}{2} - \lambda)k} \|\chi_{\{r > 2^{-k}\}} r^{-\lambda} \phi\|_{L_t^2 L_x^\infty} \\ &\quad + \sup_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi\|_{L_t^2 L_x^2} + \|(\partial_t + \partial_r) \phi\|_{Z_k^+} + \|(\partial_t - \partial_r) \phi\|_{Z_k^-}, \end{aligned}$$

where Z_k^\pm are atomic spaces defined below in Definition 3.2.

We note that the space S_k scales like free waves with $\dot{H}_x^1 \times L_x^2$ initial data. The restrictions of the spaces S and S_k to a time interval I are denoted by $S[I]$, respectively $S_k[I]$, with the induced norms. Next we provide the definition of the atomic spaces Z_k^\pm .

Definition 3.2. Let $k \in \mathbb{Z}$ and let ψ be a radially symmetric function on \mathbb{R}^{1+2} with Fourier support in $\{|\xi| \sim 2^k\}$. We introduce the auxiliary norm

$$\|\psi\|_{Y_k} := \sup_{0 < \lambda < 1} 2^{-(\frac{1}{2} + \lambda)k} \|\chi_{\{r > 2^{-k}\}} r^{-\lambda} \psi\|_{L_t^2 L_x^\infty} + \sup_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \psi\|_{L_t^2 L_x^2}.$$

Then we define

$$\begin{aligned} \|\psi\|_{Z_k^+} := \inf_{\psi = \psi^{(+)} + \psi^{(-)}} & \left\{ \left(\sup_{0 < \lambda \leq \frac{1}{2}} 2^{-(\frac{1}{2} - \lambda)k} \|r^\lambda \psi^{(+)}\|_{L_{t+r}^2 L_{t-r}^\infty} + \|\psi^{(+)}\|_{Y_k} \right) \right. \\ & \left. + \left(\sup_{2 \leq q, p \leq \infty, \frac{1}{q} + \frac{1}{p} < \frac{1}{2}} \sup_{0 < \lambda < 1} 2^{(\frac{1}{q} + \frac{2}{p} - 1)k} 2^{\lambda k} \sum_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{+\lambda} \psi^{(-)}\|_{L_t^q L_x^p} + \|\psi^{(-)}\|_{Y_k} \right) \right\} \end{aligned}$$

and

$$\begin{aligned} \|\psi\|_{Z_k^-} := \inf_{\psi = \psi^{(+)} + \psi^{(-)}} & \left\{ \left(\sup_{2 \leq q, p \leq \infty, \frac{1}{q} + \frac{1}{p} < \frac{1}{2}} \sup_{0 < \lambda < 1} 2^{(\frac{1}{q} + \frac{2}{p} - 1)k} 2^{\lambda k} \sum_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^\lambda \psi^{(+)}\|_{L_t^q L_x^p} + \|\psi^{(+)}\|_{Y_k} \right) \right. \\ & \left. + \left(\sup_{0 < \lambda \leq \frac{1}{2}} 2^{-(\frac{1}{2} - \lambda)k} \|r^{+\lambda} \psi^{(-)}\|_{L_{t-r}^2 L_{t+r}^\infty} + \|\psi^{(-)}\|_{Y_k} \right) \right\}. \end{aligned}$$

We will place the nonlinearities in the simple $L_t^1 L_x^2$ space. The remainder of this section is devoted to the proof of the following key energy estimate connecting the S_k space and the $L_t^1 L_x^2$ space.

Lemma 3.3 (Energy estimate). Let $k \in \mathbb{Z}$ and let I be any time interval containing 0. For any radially symmetric function ϕ on $I \times \mathbb{R}^2$ with Fourier support in $\{|\xi| \sim 2^k\}$, we have

$$(3.1) \quad \|\phi_k\|_{S_k[I]} \lesssim \|\nabla_{t,x} \phi_k(0)\|_{L_x^2} + \|\square \phi_k\|_{L_t^1 L_x^2[I]}.$$

The proof of Lemma 3.3 is an immediate consequence of the next lemmas. Here we first note that in the radial context a significantly larger range of admissible Strichartz norms is at our disposal, see Sterbenz [28] and Fang-Wang [5].

Lemma 3.4 (Strichartz estimates). Let $2 \leq q, p \leq \infty$ with $(q, p) \neq (\infty, \infty)$ satisfy $\frac{1}{q} + \frac{1}{p} < \frac{1}{2}$. Let $k \in \mathbb{Z}$ and let I be any time interval containing 0. For any radially symmetric function ϕ on $I \times \mathbb{R}^2$ with Fourier support in $\{|\xi| \sim 2^k\}$, it holds that

$$2^{(\frac{1}{q} + \frac{2}{p} - 1)k} \|\nabla_{t,x} \phi_k\|_{L_t^q L_x^p[I]} \lesssim \|\nabla_{t,x} \phi(0)\|_{L_x^2} + \|\square \phi\|_{L_t^1 L_x^2[I]}.$$

Next we prove Strichartz estimates involving a radial weight.

Lemma 3.5 (Weighted Strichartz estimates). Let $2 \leq p < \infty$ and $\frac{1}{p} < \lambda < \frac{2}{p}$. Let $k \in \mathbb{Z}$ and let I be any time interval containing 0. For any radially symmetric function ϕ on $I \times \mathbb{R}^2$ with Fourier support in $\{|\xi| \sim 2^k\}$, we have

$$2^{(\frac{1}{2} + \frac{2}{p} - \lambda)k} \|r^{-\lambda} \phi\|_{L_t^2 L_x^p[I]} \lesssim \|\nabla_{t,x} \phi(0)\|_{L_x^2} + \|\square \phi\|_{L_t^1 L_x^2[I]}.$$

Proof. We adapt the proof of Strichartz estimates under the spherical symmetry assumption by Sterbenz [28] to incorporate the radial weight $r^{-\lambda}$. In view of Duhamel's formula we may assume without loss of generality that ϕ is of the form $\phi = e^{it|\nabla|} f_k$ for a radially symmetric function f_k with Fourier support on $|\xi| \sim 2^k$. Moreover, by scaling invariance it suffices to consider the case $k = 0$.

First, we recall that for radially symmetric initial data, the free wave propagator in two space dimensions takes on the specific form

$$(e^{it|\nabla|} f_0)(x) = \int_{\mathbb{R}^2} e^{2\pi i(t|\xi| + x \cdot \xi)} \hat{f}_0(\xi) d\xi = 2\pi \int_0^\infty e^{2\pi i t \rho} J_0(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho d\rho,$$

where $\varphi_0(\rho)$ is a smooth bump function supported on $\{\rho \sim 1\}$ and where

$$J_0(y) = \frac{1}{2\pi} \int_0^{2\pi} e^{iy \sin(\theta)} d\theta, \quad y \in \mathbb{R},$$

is the Bessel function of order 0. Moreover, we recall the following standard asymptotics for the Bessel function of order 0, see e.g. [40],

$$(3.2) \quad J_0(y) = y^{-\frac{1}{2}} e^{+iy} \beta_+(y) + y^{-\frac{1}{2}} e^{-iy} \beta_-(y), \quad y \geq 1,$$

where the functions β_{\pm} satisfy the symbol-type bounds

$$(3.3) \quad |\beta_{\pm}^{(n)}(y)| \lesssim_n y^{-n}, \quad n \geq 0, y \geq 1.$$

We now distinguish the two regimes $r \lesssim 1$ and $r \gg 1$. In the former case, we just use Hölder's inequality, the assumption $\lambda < \frac{2}{p}$ and a standard TT^* estimate to obtain that

$$\begin{aligned} \|\chi_{\{r \lesssim 1\}} r^{-\lambda} e^{it|\nabla|} f_0\|_{L_t^2 L_x^p} &\lesssim \|\chi_{\{r \lesssim 1\}} r^{-\lambda}\|_{L_x^p} \|\chi_{\{r \lesssim 1\}} \int_0^\infty e^{2\pi i t \rho} J_0(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho d\rho\|_{L_t^2} \\ &\lesssim \|\hat{f}_0(\rho) \rho\|_{L_p^2} \\ &\lesssim \|f_0\|_{L_x^2}. \end{aligned}$$

In the latter case $r \gg 1$, the asymptotics (3.2) yield that

$$\chi_{\{r \gg 1\}} r^{-\lambda} (e^{it|\nabla|} f_0)(r) \sim \sum_{\pm} \chi_{\{r \gg 1\}} r^{-\frac{1}{2}-\lambda} \int_0^\infty e^{2\pi i(t \pm r)\rho} \beta_{\pm}(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho^{\frac{1}{2}} d\rho.$$

Noting that the function $\rho \mapsto \hat{f}_0(\rho) \rho^{\frac{1}{2}}$ is compactly supported in the interval $(0, 4)$, we may consider its Fourier series

$$\hat{f}_0(\rho) \rho^{\frac{1}{2}} = \sum_{n \in \mathbb{Z}} c_n e^{i\frac{\pi}{2} n \rho}, \quad \rho \in (0, 4),$$

whose Fourier coefficients satisfy

$$\sum_{n \in \mathbb{Z}} |c_n|^2 \sim \|\hat{f}_0(\rho) \rho^{\frac{1}{2}}\|_{L_p^2}^2 \sim \|f_0\|_{L_x^2}^2.$$

We can therefore write

$$\chi_{\{r \gg 1\}} r^{-\lambda} (e^{it|\nabla|} f_0)(r) \sim \sum_{\pm} r^{-\frac{1}{2}-\lambda} \sum_{n \in \mathbb{Z}} c_n \psi_n^{\pm}(t, r)$$

with

$$\psi_n^{\pm}(t, r) = \chi_{\{r \gg 1\}} \int_0^\infty e^{2\pi i(t \pm r + \frac{n}{4})\rho} \beta_{\pm}(2\pi r \rho) \varphi_0(\rho) d\rho.$$

From the symbol-type bounds (3.3) we obtain by repeated integration by parts that

$$|\psi_n(t, r)| \lesssim_M \left(1 + \left|t \pm r + \frac{n}{4}\right|\right)^{-M}.$$

Thus, choosing $M \gg 1$ sufficiently large, we infer from an application of Hölder's inequality (in $n \in \mathbb{Z}$) and the embedding $\ell^p \hookrightarrow \ell^2$ for $p \geq 2$ that

$$\begin{aligned}
\|\chi_{\{r \gg 1\}} r^{-\lambda} e^{it|\nabla|} f_0\|_{L_x^p} &\lesssim \sum_{\pm} \left(\int_1^\infty \left(\sum_{n \in \mathbb{Z}} |c_n| |\psi_n^\pm(t, r)| \right)^p r^{1-p(\frac{1}{2}+\lambda)} dr \right)^{\frac{1}{p}} \\
&\lesssim \sum_{\pm} \left(\int_1^\infty \left(\sum_{n \in \mathbb{Z}} \frac{|c_n|}{(1 + |t \pm r + \frac{n}{4}|)^M} \right)^p r^{1-p(\frac{1}{2}+\lambda)} dr \right)^{\frac{1}{p}} \\
&\lesssim \left(\int_1^\infty \left(\sum_{n \in \mathbb{Z}} \frac{|c_n|^p}{(1 + |t + \frac{n}{4} - r|)^{(M-2)p}} \right) r^{1-p(\frac{1}{2}+\lambda)} dr \right)^{\frac{1}{p}} \\
&\lesssim \left(\sum_{n \in \mathbb{Z}} \frac{|c_n|^p}{(1 + |t + \frac{n}{4}|)^{p(\frac{1}{2}+\lambda)-1}} \right)^{\frac{1}{p}} \\
&\lesssim \left(\sum_{n \in \mathbb{Z}} \frac{|c_n|^2}{(1 + |t + \frac{n}{4}|)^{2(\frac{1}{2}+\lambda-\frac{1}{p})}} \right)^{\frac{1}{2}}.
\end{aligned}$$

Since by assumption $\lambda > \frac{1}{p}$, we obtain the desired estimate

$$\|\chi_{\{r \gg 1\}} r^{-\lambda} e^{it|\nabla|} f_0\|_{L_t^2 L_x^p} \lesssim \left(\sum_{n \in \mathbb{Z}} |c_n|^2 \right)^{\frac{1}{2}} \sim \|f_0\|_{L_x^2}.$$

□

Now we are in the position to deduce a weighted $L_t^2 L_x^\infty$ Strichartz estimate.

Lemma 3.6 (Weighted endpoint Strichartz estimate). *Let $0 < \lambda < 1$. Let $k \in \mathbb{Z}$ and let I be any time interval containing 0. For any radially symmetric function ϕ on $I \times \mathbb{R}^2$ with Fourier support in $\{|\xi| \sim 2^k\}$, we have*

$$2^{(\frac{1}{2}-\lambda)k} \|\chi_{\{r \geq 2^{-k}\}} r^{-\lambda} \phi\|_{L_t^2 L_x^\infty[I]} \lesssim \|\nabla_{t,x} \phi(0)\|_{L_x^2} + \|\square \phi\|_{L_t^1 L_x^2[I]}.$$

Proof. We begin by writing

$$\|\chi_{\{r \geq 2^{-k}\}} r^{-\lambda} \phi\|_{L_t^2 L_x^\infty} \lesssim \sum_{\ell \geq -k} \|\chi_{\{2^\ell \leq r \leq 2^{\ell+1}\}} r^{-\lambda} P_k \phi\|_{L_t^2 L_x^\infty}.$$

Since the Littlewood-Paley projection operator P_k lives at spatial scale 2^{-k} , for $\ell \geq -k$ it approximately preserves the spatial localization enforced by the cutoff $\chi_{\{2^\ell \leq r \leq 2^{\ell+1}\}}$ up to exponentially decaying tails that can be dealt with easily. We may therefore replace the right-hand side of the previous line by

$$\sum_{\ell \geq -k} \|\chi_{\{2^\ell \leq r \leq 2^{\ell+1}\}} r^{-\lambda} \tilde{P}_k(\chi_{\{2^{\ell-2} \leq r \leq 2^{\ell+3}\}} P_k \phi)\|_{L_t^2 L_x^\infty},$$

where \tilde{P}_k is a fattened Littlewood-Paley projection operator such that $\tilde{P}_k P_k \phi = P_k \phi$. Applying Bernstein's inequality for some $2 \leq p < \infty$ with $\frac{1}{p} < \frac{\lambda}{2} < \frac{2}{p}$, we arrive at the bound

$$\sum_{\ell \geq -k} 2^{-\lambda \ell} 2^{\frac{2}{p} k} \|\chi_{\{2^{\ell-2} \leq r \leq 2^{\ell+3}\}} P_k \phi\|_{L_t^2 L_x^p} \lesssim \sum_{\ell \geq -k} 2^{-\frac{\lambda}{2} \ell} 2^{\frac{2}{p} k} \|r^{-\frac{\lambda}{2}} P_k \phi\|_{L_t^2 L_x^p}.$$

Next we invoke the weighted Strichartz estimates from Lemma 3.5 to obtain the bound

$$\sum_{\ell \geq -k} 2^{-\frac{\lambda}{2} \ell} 2^{\frac{2}{p} k} 2^{-(\frac{1}{2} + \frac{2}{p} - \frac{\lambda}{2})k} (\|\nabla_{t,x} \phi(0)\|_{L_x^2} + \|\square \phi\|_{L_t^1 L_x^2}) \lesssim 2^{-(\frac{1}{2}-\lambda)k} (\|\nabla_{t,x} \phi(0)\|_{L_x^2} + \|\square \phi\|_{L_t^1 L_x^2}),$$

which finishes the proof. □

Moreover, we have the following local energy decay estimate.

Lemma 3.7 (Local energy decay estimate). *Let $k \in \mathbb{Z}$ and let I be any time interval containing 0. For any radially symmetric function ϕ on $I \times \mathbb{R}^2$ with Fourier support in $\{|\xi| \sim 2^k\}$, it holds that*

$$\sup_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi\|_{L_t^2 L_x^2[I]} \lesssim \|\nabla_{t,x} \phi(0)\|_{L_x^2} + \|\square \phi\|_{L_t^1 L_x^2[I]}.$$

Proof. We proceed similarly to the proof of Lemma 3.5. In view of Duhamel's formula, we may again assume that ϕ is of the form $\phi = e^{it|\nabla|} f_k$ for a radially symmetric function f_k with Fourier support on $|\xi| \sim 2^k$. Moreover, by scaling invariance it suffices to consider the case $k = 0$.

We first prove the local energy decay estimate for the time derivative $\partial_t(e^{it|\nabla|} f_0)$. Due to the radial symmetry assumption, we have the specific representation formula

$$\partial_t(e^{it|\nabla|} f_0)(r) = 4\pi^2 i \int_0^\infty e^{2\pi i t \rho} J_0(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho^2 d\rho.$$

We distinguish the regimes $r \lesssim 1$ and $r \gg 1$. Here we only treat the more difficult case $r \gg 1$. By the asymptotics (3.2) we have that

$$\chi_{\{r \gg 1\}} r^{-\frac{1}{2}} \partial_t(e^{it|\nabla|} f_0)(r) \sim \sum_{\pm} \chi_{\{r \gg 1\}} r^{-1} \int_0^\infty e^{2\pi i(t \pm r)\rho} \beta_{\pm}(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho^{\frac{3}{2}} d\rho.$$

The function $\rho \mapsto \hat{f}_0(\rho) \rho^{\frac{3}{2}}$ has compact support in the interval $(0, 4)$ and can therefore be developed into a Fourier series

$$\hat{f}_0(\rho) \rho^{\frac{3}{2}} = \sum_{n \in \mathbb{Z}} c_n e^{i\frac{\pi}{2} n \rho}, \quad \rho \in (0, 4),$$

where the Fourier coefficients satisfy

$$\sum_{n \in \mathbb{Z}} |c_n|^2 \sim \|\hat{f}_0(\rho) \rho^{\frac{3}{2}}\|_{L_\rho^2}^2 \lesssim \|f_0\|_{L_x^2}^2.$$

Now we follow closely the arguments in the proof of Lemma 3.5 to find that

$$\sup_{\ell \geq 0} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \partial_t(e^{it|\nabla|} f_0)(r)\|_{L_t^2 L_x^2} \lesssim \sup_{\ell \geq 0} \left(\sum_{\pm} \int_{\mathbb{R}} \int_{\{r \sim 2^\ell\}} \sum_{n \in \mathbb{Z}} \frac{|c_n|^2}{(1 + |t \pm r + \frac{n}{4}|)^2} r^{-1} dr dt \right)^{\frac{1}{2}}.$$

Changing the order of integration and computing the time integral first, we see that the right-hand side obeys the desired bound

$$\sup_{\ell \geq 0} \left(\int_{\{r \sim 2^\ell\}} \sum_{n \in \mathbb{Z}} |c_n|^2 r^{-1} dr \right)^{\frac{1}{2}} \lesssim \left(\sum_{n \in \mathbb{Z}} |c_n|^2 \right)^{\frac{1}{2}} \lesssim \|f_0\|_{L_x^2}.$$

The proof of the local energy decay estimate for the spatial derivatives of $e^{it|\nabla|} f_0$ proceeds analogously by noting that the first derivative J'_0 of the Bessel function J_0 of order 0 satisfies the same asymptotics (3.2) as J_0 . \square

Finally, we turn to the atomic Z_k^\pm spaces.

Lemma 3.8 (The Z_k^\pm spaces). *Let $k \in \mathbb{Z}$ and let I be any time interval containing 0. For any radially symmetric function ϕ on $I \times \mathbb{R}^2$ with Fourier support in $\{|\xi| \sim 2^k\}$, we have that*

$$\|(\partial_t + \partial_r) \phi_k\|_{Z_k^+[I]} \lesssim \|\nabla_{t,x} \phi_k(0)\|_{L_x^2} + \|\square \phi_k\|_{L_t^1 L_x^2[I]}$$

and

$$\|(\partial_t - \partial_r) \phi_k\|_{Z_k^-[I]} \lesssim \|\nabla_{t,x} \phi_k(0)\|_{L_x^2} + \|\square \phi_k\|_{L_t^1 L_x^2[I]}.$$

Proof. We only provide the proof of the estimate for the Z_k^+ space since the case of the Z_k^- space can be dealt with analogously. Moreover, in view of Duhamel's formula we may assume that ϕ_k is of the form $e^{it|\nabla|} f_k$ for a radially symmetric function f_k with Fourier support on $\{|\xi| \sim 2^k\}$. By scaling invariance it suffices to consider the case $k = 0$.

Then we observe that

$$(\partial_t + \partial_r)(e^{it|\nabla|} f_0)(r) = 4\pi^2 \int_0^\infty e^{2\pi i t \rho} (iJ_0(2\pi r \rho) + J_0'(2\pi r \rho)) \varphi_0(\rho) \hat{f}_0(\rho) \rho^2 d\rho.$$

By stationary phase it is easy to see that we have the asymptotics

$$(3.4) \quad iJ_0(y) + J_0'(y) = \frac{i}{2\pi} \int_0^{2\pi} e^{iy \sin(\theta)} (1 + \sin(\theta)) d\theta = y^{-\frac{1}{2}} e^{+iy} \beta_+(y) + y^{-\frac{3}{2}} e^{-iy} \beta_-(y), \quad y \geq 1,$$

where the functions β_\pm satisfy the symbol-type bounds

$$|\beta_\pm^{(n)}(y)| \lesssim_n y^{-n}, \quad n \geq 0, y \geq 1.$$

We now distinguish the regimes $r \lesssim 1$ and $r \gg 1$. Here we only turn to the more delicate latter case $r \gg 1$. By the asymptotics (3.4) we see that $\chi_{\{r \gg 1\}}(\partial_t + \partial_r)(e^{it|\nabla|} f_0)$ decomposes into two components

$$(3.5) \quad \begin{aligned} \chi_{\{r \gg 1\}}(\partial_t + \partial_r)(e^{it|\nabla|} f_0)(r) &\sim \chi_{\{r \gg 1\}} r^{-\frac{1}{2}} \int_0^\infty e^{2\pi i(t+r)\rho} \beta_+(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho^{\frac{3}{2}} d\rho \\ &\quad + \chi_{\{r \gg 1\}} r^{-\frac{3}{2}} \int_0^\infty e^{2\pi i(t-r)\rho} \beta_-(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho^{\frac{1}{2}} d\rho. \end{aligned}$$

It is easy to see that both components on the right-hand side have finite Y_0 norm. Moreover, the first component on the right-hand side of (3.5) can be placed into the weighted $L_{t+r}^2 L_{t-r}^\infty$ component of the Z_0^+ space. More precisely, for any $0 < \lambda \leq \frac{1}{2}$ we obtain by a standard TT^* argument that

$$\begin{aligned} &\left\| r^{+\lambda} \chi_{\{r \gg 1\}} r^{-\frac{1}{2}} \int_0^\infty e^{2\pi i(t+r)\rho} \beta_+(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho^{\frac{3}{2}} d\rho \right\|_{L_{t+r}^2 L_{t-r}^\infty} \\ &\lesssim \left\| \int_0^\infty e^{2\pi i(t+r)\rho} \chi_{\{r \gg 1\}} \beta_+(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho^{\frac{3}{2}} d\rho \right\|_{L_{t+r}^2 L_{t-r}^\infty} \\ &\lesssim \|\hat{f}_0(\rho) \rho^{\frac{3}{2}}\|_{L_\rho^2} \\ &\lesssim \|f_0\|_{L_x^2}. \end{aligned}$$

On the other hand, the second component on the right-hand side of (3.5) satisfies for any exponent pair (q, p) with $2 \leq q, p \leq \infty$ and $\frac{1}{q} + \frac{1}{p} < \frac{1}{2}$ and for any $0 < \lambda < 1$ that

$$\begin{aligned} &\sum_{\ell \in \mathbb{Z}} \left\| \chi_{\{r \sim 2^\ell\}} r^{+\lambda} \chi_{\{r \gg 1\}} r^{-\frac{3}{2}} \int_0^\infty e^{2\pi i(t-r)\rho} \beta_-(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho^{\frac{1}{2}} d\rho \right\|_{L_t^q L_x^p} \\ &\lesssim \sum_{\ell \gg 1} 2^{(\lambda-1)\ell} \left\| r^{-\frac{1}{2}} \int_0^\infty e^{2\pi i(t-r)\rho} \chi_{\{r \gg 1\}} \beta_-(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho^{\frac{1}{2}} d\rho \right\|_{L_t^q L_x^p} \\ &\lesssim \left\| r^{-\frac{1}{2}} \int_0^\infty e^{2\pi i(t-r)\rho} \chi_{\{r \gg 1\}} \beta_-(2\pi r \rho) \varphi_0(\rho) \hat{f}_0(\rho) \rho^{\frac{1}{2}} d\rho \right\|_{L_t^q L_x^p} \\ &\lesssim \|f_0\|_{L_x^2}, \end{aligned}$$

where in the case of $(q, p) = (\infty, \infty)$ the last estimate just follows from Hölder's inequality while for all other admissible exponent pairs (q, p) the last estimate follows from the proof of Strichartz estimates in the radial case as in Sterbenz [28]. Putting things together, we conclude that

$$\|(\partial_t + \partial_r)(e^{it|\nabla|} f_0)\|_{Z_0^+} \lesssim \|f_0\|_{L_x^2},$$

which finishes the proof of Lemma 3.8. \square

4. MULTILINEAR ESTIMATES

Here we collect several important multilinear estimates that will be of crucial use to estimate the wave maps nonlinearity. We begin with the following null form estimate.

Proposition 4.1 (Null form estimate). *For $\frac{3}{2} \leq p \leq \infty$ it holds that*

$$(4.1) \quad \left\| \partial_\alpha \phi_{k_1}^{(1)} \partial^\alpha \phi_{k_2}^{(2)} \right\|_{L_t^p L_x^p} \lesssim 2^{(1-\frac{3}{2p})k_1} 2^{(1-\frac{3}{2p})k_2} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}}.$$

Proof. The assertion follows by interpolation between the simple $L_t^\infty L_x^\infty$ estimate

$$\left\| \partial_\alpha \phi_{k_1}^{(1)} \partial^\alpha \phi_{k_2}^{(2)} \right\|_{L_t^\infty L_x^\infty} \lesssim 2^{k_1} \left\| \nabla_{t,x} \phi_{k_1} \right\|_{L_t^\infty L_x^2} 2^{k_2} \left\| \nabla_{t,x} \phi_{k_2} \right\|_{L_t^\infty L_x^2} \lesssim 2^{k_1} 2^{k_2} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}}$$

and the following $L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}$ estimate

$$(4.2) \quad \left\| \partial_\alpha \phi_{k_1}^{(1)} \partial^\alpha \phi_{k_2}^{(2)} \right\|_{L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}} \lesssim \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}},$$

which we now prove. Since in the radially symmetric setting it holds that

$$-2\partial_\alpha \phi \partial^\alpha \psi = (\partial_t + \partial_r)\phi(\partial_t - \partial_r)\psi + (\partial_t - \partial_r)\phi(\partial_t + \partial_r)\psi,$$

by symmetry it suffices to consider bounding the expression $(\partial_t + \partial_r)\phi_{k_1}^{(1)}(\partial_t - \partial_r)\phi_{k_2}^{(2)}$ in $L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}$. To this end we decompose the two inputs into (+) and (-) components, i.e.

$$(\partial_t + \partial_r)\phi_{k_1}^{(1)} = (\partial_t + \partial_r)\phi_{k_1}^{(1,+)} + (\partial_t + \partial_r)\phi_{k_1}^{(1,-)}$$

and

$$(\partial_t - \partial_r)\phi_{k_2}^{(2)} = (\partial_t - \partial_r)\phi_{k_2}^{(2,+)} + (\partial_t - \partial_r)\phi_{k_2}^{(2,-)}.$$

Then we consider all possible interactions. In the case of (+)/(+) interactions we estimate

$$\begin{aligned} & \left\| (\partial_t + \partial_r)\phi_{k_1}^{(1,+)} (\partial_t - \partial_r)\phi_{k_2}^{(2,+)} \right\|_{L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}} \\ & \lesssim \sup_{\ell \in \mathbb{Z}} \left\| \chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} (\partial_t + \partial_r)\phi_{k_1}^{(1,+)} \right\|_{L_t^2 L_x^2} \sum_{\ell \in \mathbb{Z}} \left\| \chi_{\{r \sim 2^\ell\}} r^{+\frac{1}{2}} (\partial_t - \partial_r)\phi_{k_2}^{(2,+)} \right\|_{L_t^6 L_x^6} \\ & \lesssim \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}} \end{aligned}$$

and the case of (-)/(+) as well as of (-)/(-) interactions can be bounded analogously. We are therefore left with the treatment of the delicate (+)/(-) interactions. By spherical symmetry, we may write

$$\left\| (\partial_t + \partial_r)\phi_{k_1}^{(1,+)} (\partial_t - \partial_r)\phi_{k_2}^{(2,-)} \right\|_{L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}}^{\frac{3}{2}} \lesssim \iint r^{-\frac{1}{2}} \left| r^{\frac{1}{2}} (\partial_t + \partial_r)\phi_{k_1}^{(1,+)} \right|^{\frac{3}{2}} \left| r^{\frac{1}{2}} (\partial_t - \partial_r)\phi_{k_2}^{(2,-)} \right|^{\frac{3}{2}} dr dt.$$

Changing variables from (t, r) to $(t+r, t-r)$ and noting that $r = \frac{1}{2}((t+r) - (t-r))$, we can estimate the previous line by

$$\iint |(t+r) - (t-r)|^{-\frac{1}{2}} \left(\left\| r^{\frac{1}{2}} (\partial_t + \partial_r)\phi_{k_1}^{(1,+)} \right\|_{L_{t-r}^\infty}^{\frac{3}{2}} \right) (t+r) \left(\left\| r^{\frac{1}{2}} (\partial_t - \partial_r)\phi_{k_2}^{(2,-)} \right\|_{L_{t+r}^\infty}^{\frac{3}{2}} \right) (t-r) d(t+r) d(t-r).$$

Next we apply Hölder's inequality followed by the Hardy-Littlewood-Sobolev inequality and obtain the desired bound

$$\begin{aligned}
& \left\| \left\| r^{\frac{1}{2}}(\partial_t + \partial_r)\phi_{k_1}^{(1,+)} \right\|_{L_{t-r}^\infty}^{\frac{3}{2}} \left\| \int |(t+r) - (t-r)|^{-\frac{1}{2}} \left(\left\| r^{\frac{1}{2}}(\partial_t - \partial_r)\phi_{k_2}^{(2,-)} \right\|_{L_{t+r}^\infty}^{\frac{3}{2}} \right) (t-r) d(t-r) \right\|_{L_{t+r}^4} \right. \\
& \lesssim \left\| r^{\frac{1}{2}}(\partial_t + \partial_r)\phi_{k_1}^{(1,+)} \right\|_{L_{t+r}^2 L_{t-r}^\infty}^{\frac{3}{2}} \left\| r^{\frac{1}{2}}(\partial_t - \partial_r)\phi_{k_2}^{(2,-)} \right\|_{L_{t+r}^\infty}^{\frac{3}{2}} \left\| \right\|_{L_{t-r}^4} \\
& \lesssim \left\| r^{\frac{1}{2}}(\partial_t + \partial_r)\phi_{k_1}^{(1,+)} \right\|_{L_{t+r}^2 L_{t-r}^\infty}^{\frac{3}{2}} \left\| r^{\frac{1}{2}}(\partial_t - \partial_r)\phi_{k_2}^{(2,-)} \right\|_{L_{t-r}^2 L_{t+r}^\infty}^{\frac{3}{2}} \\
& \lesssim \|\phi_{k_1}^{(1)}\|_{S_{k_1}}^{\frac{3}{2}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}}^{\frac{3}{2}}.
\end{aligned}$$

This finishes the proof of the $L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}$ estimate (4.2) and thus concludes the proof of Proposition 4.1. \square

We emphasize that the preceding proposition allows to estimate the null form $\partial_\alpha \phi \partial^\alpha \psi$ “below” $L_t^2 L_x^2$, in fact in $L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}$. This feature will be crucial in the concentration compactness step in Section 7 to gain smallness for certain multilinear expressions, which is explained in detail in Subsection 7.4.3. It was first pointed out by Klainerman-Machedon [8] that the null form $\partial_\alpha \phi \partial^\alpha \psi$ can be estimated “below” $L_t^2 L_x^2$ provided both factors are free radial waves. The corresponding improvements without the radially assumption are due to Bourgain [2], Wolff [41] and Tao [35]. Next we establish a weighted null form estimate.

Proposition 4.2 (Weighted null form estimate). *For $0 < \lambda \leq \frac{1}{2}$ it holds that*

$$(4.3) \quad \left\| r^{+\lambda} \partial_\alpha \phi_{k_1}^{(1)} \partial^\alpha \phi_{k_2}^{(2)} \right\|_{L_t^2 L_x^2} \lesssim 2^{(\frac{1}{2}-\lambda) \min\{k_1, k_2\}} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}}.$$

Proof. Without loss of generality we may assume that $k_1 \leq k_2$. Then we can dispense with the case when the radial variable is restricted to the range $r \leq 2^{-k_1}$ because here we can just bound by

$$\begin{aligned}
\left\| \chi_{\{r \leq 2^{-k_1}\}} r^{+\lambda} \partial_\alpha \phi_{k_1}^{(1)} \partial^\alpha \phi_{k_2}^{(2)} \right\|_{L_t^2 L_x^2} & \lesssim \sum_{\ell \leq -k_1} \left\| \chi_{\{r \sim 2^\ell\}} r^{\frac{1}{2}+\lambda} \partial_\alpha \phi_{k_1}^{(1)} \right\|_{L_t^\infty L_x^\infty} \left\| \chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \partial^\alpha \phi_{k_2}^{(2)} \right\|_{L_t^2 L_x^2} \\
& \lesssim \sum_{\ell \leq -k_1} 2^{(\frac{1}{2}+\lambda)\ell} 2^{k_1} \|\nabla_{t,x} \phi_{k_1}^{(1)}\|_{L_t^\infty L_x^2} \sup_{\ell \in \mathbb{Z}} \left\| \chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_{k_2}^{(2)} \right\|_{L_t^2 L_x^2} \\
& \lesssim 2^{(\frac{1}{2}-\lambda)k_1} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}}.
\end{aligned}$$

Moreover, since in the radially symmetric setting the null form $\partial_\alpha \phi_{k_1}^{(1)} \partial^\alpha \phi_{k_2}^{(2)}$ can be written as a linear combination of terms $(\partial_t \pm \partial_r)\phi_{k_1}^{(1)}(\partial_t \mp \partial_r)\phi_{k_2}^{(2)}$, we are left to estimate

$$\left\| \chi_{\{r > 2^{-k_1}\}} r^{+\lambda} (\partial_t + \partial_r)\phi_{k_1}^{(1)} (\partial_t - \partial_r)\phi_{k_2}^{(2)} \right\|_{L_t^2 L_x^2}.$$

To this end we again decompose the two inputs into (+) and (−) components

$$(\partial_t \pm \partial_r)\phi_{k_j}^{(j)} = (\partial_t \pm \partial_r)\phi_{k_j}^{(j,+)} + (\partial_t \pm \partial_r)\phi_{k_j}^{(j,-)}, \quad j = 1, 2$$

and then estimate all possible interactions. In the case of (+)/(−) interactions, we have that

$$\begin{aligned}
\left\| \chi_{\{r > 2^{-k_1}\}} r^{+\lambda} (\partial_t + \partial_r)\phi_{k_1}^{(1,+)} (\partial_t - \partial_r)\phi_{k_2}^{(2,-)} \right\|_{L_t^2 L_x^2} & \lesssim \left\| r^{+\lambda} (\partial_t + \partial_r)\phi_{k_1}^{(1,+)} \right\|_{L_{t+r}^2 L_{t-r}^\infty} \left\| r^{+\frac{1}{2}} (\partial_t - \partial_r)\phi_{k_2}^{(2,-)} \right\|_{L_{t-r}^2 L_{t+r}^\infty} \\
& \lesssim 2^{(\frac{1}{2}-\lambda)k_1} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}}.
\end{aligned}$$

Then we bound the (+)/(+) interactions by

$$\begin{aligned}
& \|\chi_{\{r>2^{-k_1}\}} r^{+\lambda} (\partial_t + \partial_r) \phi_{k_1}^{(1,+)} (\partial_t - \partial_r) \phi_{k_2}^{(2,-)}\|_{L_t^2 L_x^2} \\
& \lesssim \|\chi_{\{r>2^{-k_1}\}} r^{-\lambda} (\partial_t + \partial_r) \phi_{k_1}^{(+)}\|_{L_t^2 L_x^\infty} \sum_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{+2\lambda} (\partial_t - \partial_r) \phi_{k_2}^{(+)}\|_{L_t^\infty L_x^2} \\
& \lesssim 2^{(\frac{1}{2}+\lambda)k_1} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} 2^{-2\lambda k_2} \|\phi_{k_2}^{(2)}\|_{S_{k_2}} \\
& \lesssim 2^{(\frac{1}{2}-\lambda)k_1} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}}.
\end{aligned}$$

In the case of (-)/(-) interactions we estimate as follows

$$\begin{aligned}
& \|\chi_{\{r>2^{-k_1}\}} r^{+\lambda} (\partial_t + \partial_r) \phi_{k_1}^{(1,-)} (\partial_t - \partial_r) \phi_{k_2}^{(2,-)}\|_{L_t^2 L_x^2} \\
& \lesssim \sum_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{\frac{1}{2}+\lambda} (\partial_t + \partial_r) \phi_{k_1}^{(1,-)}\|_{L_t^\infty L_x^\infty} \sup_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} (\partial_t - \partial_r) \phi_{k_2}^{(2,+)}\|_{L_t^2 L_x^2} \\
& \lesssim 2^{(\frac{1}{2}-\lambda)k_1} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}}
\end{aligned}$$

and the remaining case of (-)/(+) interactions can be treated in exactly the same manner. \square

We conclude with the following delicate trilinear estimate.

Proposition 4.3 (Trilinear estimate). *We have that*

$$(4.4) \quad \|L(\phi_{k_1}^{(1)}, \partial_\alpha \phi_{k_2}^{(2)}, \partial^\alpha \phi_{k_3}^{(3)})\|_{L_t^1 L_x^2} \lesssim 2^{-\frac{1}{4}(k_1 - \min\{k_2, k_3\})} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}} \|\phi_{k_3}^{(3)}\|_{S_{k_3}}$$

for $k_1 \geq \min\{k_2, k_3\} + O(1)$.

Proof. We may assume without loss of generality that $k_2 \leq k_3$. Moreover, we may restrict the radial variable to the range $\{r > 2^{-k_2}\}$, because otherwise we can just easily estimate

$$\begin{aligned}
& \|\chi_{\{r \leq 2^{-k_2}\}} L(\phi_{k_1}^{(1)}, \partial_\alpha \phi_{k_2}^{(2)}, \partial^\alpha \phi_{k_3}^{(3)})\|_{L_t^1 L_x^2} \\
& \lesssim \|\phi_{k_1}^{(1)}\|_{L_t^4 L_x^\infty} \sum_{\ell \leq -k_2} \|\chi_{\{r \sim 2^\ell\}} r^{+\frac{1}{2}} \nabla_{t,x} \phi_{k_2}^{(2)}\|_{L_t^4 L_x^\infty} \sup_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_{k_3}^{(3)}\|_{L_t^2 L_x^2} \\
& \lesssim 2^{-\frac{1}{4}(k_1 - k_2)} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}} \|\phi_{k_3}^{(3)}\|_{S_{k_3}}.
\end{aligned}$$

In the radially symmetric setting we have the identity

$$-2\partial_\alpha \phi \partial^\alpha \psi = (\partial_t + \partial_r) \phi (\partial_t - \partial_r) \psi + (\partial_t - \partial_r) \phi (\partial_t + \partial_r) \psi.$$

By symmetry it therefore suffices to now bound the expression

$$\|\chi_{\{r>2^{-k_2}\}} L(\phi_{k_1}^{(1)}, (\partial_t + \partial_r) \phi_{k_2}^{(2)}, (\partial_t - \partial_r) \phi_{k_3}^{(3)})\|_{L_t^1 L_x^2}.$$

To this end we decompose $(\partial_t + \partial_r) \phi_{k_2}^{(2)}$ and $(\partial_t - \partial_r) \phi_{k_3}^{(3)}$ into (+) and (-) components

$$(\partial_t \pm \partial_r) \phi_{k_j}^{(j)} = (\partial_t \pm \partial_r) \phi_{k_j}^{(j,+)} + (\partial_t \pm \partial_r) \phi_{k_j}^{(j,-)} \quad \text{for } j = 2, 3$$

and distinguish all possible interaction scenarios. In the case of (+)/(-) interactions, we estimate

$$\begin{aligned}
& \|\chi_{\{r>2^{-k_2}\}} L(\phi_{k_1}^{(1)}, (\partial_t + \partial_r) \phi_{k_2}^{(2,+)} (\partial_t - \partial_r) \phi_{k_3}^{(3,-)})\|_{L_t^1 L_x^2} \\
& \lesssim \|\chi_{\{r>2^{-k_1}\}} r^{-\frac{1}{4}} \phi_{k_1}^{(1)}\|_{L_t^2 L_x^\infty} \|r^{+\frac{1}{4}} (\partial_t + \partial_r) \phi_{k_2}^{(2,+)}\|_{L_{t+r}^2 L_{t-r}^\infty} \|r^{+\frac{1}{2}} (\partial_t - \partial_r) \phi_{k_3}^{(3,-)}\|_{L_{t-r}^2 L_{t+r}^\infty} \\
& \lesssim 2^{-\frac{1}{4}(k_1 - k_2)} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}} \|\phi_{k_3}^{(3)}\|_{S_{k_3}},
\end{aligned}$$

where we used that $k_1 \geq k_2 + O(1)$ by assumption. Then we bound the $(-)/(+)$ interactions by

$$\begin{aligned} & \|\chi_{\{r>2^{-k_2}\}} L(\phi_{k_1}^{(1)}, (\partial_t + \partial_r)\phi_{k_2}^{(2,-)}, (\partial_t - \partial_r)\phi_{k_3}^{(3,+)})\|_{L_t^1 L_x^2} \\ & \lesssim \|\chi_{\{r>2^{-k_1}\}} r^{-\frac{1}{4}} \phi_{k_1}^{(1)}\|_{L_t^2 L_x^\infty} \sum_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{+\frac{3}{4}} (\partial_t + \partial_r)\phi_{k_2}^{(2,-)}\|_{L_t^\infty L_x^\infty} \sup_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_{k_3}^{(3,+)}\|_{L_t^2 L_x^2} \\ & \lesssim 2^{-\frac{1}{4}(k_1 - k_2)} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}} \|\phi_{k_3}^{(3)}\|_{S_{k_3}} \end{aligned}$$

and the case of $(-)/(-)$ interactions can be treated in exactly the same manner. Finally, we have to consider the case of $(+)/(+)$ interactions

$$\begin{aligned} & \|\chi_{\{r>2^{-k_2}\}} L(\phi_{k_1}^{(1)}, (\partial_t + \partial_r)\phi_{k_2}^{(2,+)}, (\partial_t - \partial_r)\phi_{k_3}^{(3,+)})\|_{L_t^1 L_x^2} \\ & \lesssim \|\chi_{\{r>2^{-k_1}\}} r^{-\frac{1}{4}} \phi_{k_1}^{(1)}\|_{L_t^2 L_x^\infty} \|\chi_{\{r>2^{-k_2}\}} r^{-\frac{1}{4}} (\partial_t + \partial_r)\phi_{k_2}^{(2,+)}\|_{L_t^2 L_x^\infty} \sum_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{+\frac{1}{2}} (\partial_t - \partial_r)\phi_{k_3}^{(3,+)}\|_{L_t^\infty L_x^2} \\ & \lesssim 2^{-\frac{1}{4}(k_1 - k_2)} 2^{-\frac{1}{2}(k_3 - k_2)} \|\phi_{k_1}^{(1)}\|_{S_{k_1}} \|\phi_{k_2}^{(2)}\|_{S_{k_2}} \|\phi_{k_3}^{(3)}\|_{S_{k_3}}, \end{aligned}$$

which finishes the proof. \square

5. DECOMPOSITION OF THE NONLINEARITY AND RENORMALIZATION

In this section we study the structure of the nonlinearity in the frequency localized wave maps equation

$$(5.1) \quad \square \phi_k = -P_k(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi), \quad k \in \mathbb{Z}.$$

Our main tools to estimate this wave maps nonlinearity are the null form estimate (4.1) and the trilinear estimate (4.4) from the previous section. However, these will turn out to be far from sufficient, even for small energies. Accordingly, using Littlewood-Paley theory we will “peel off” the “good” parts from the nonlinearity to isolate its non-perturbative parts. Then we will introduce the renormalization procedure to deal with the latter parts of the nonlinearity. For the sake of readability we decided to only introduce a simplified version of our renormalization procedure in this section. This version suffices for small energies, but in order to handle large energies the construction of our renormalization procedure will, roughly speaking, involve further frequency and spatial truncations. We will explain the necessary modifications as we go in the later sections of this paper.

Our starting point is the following decomposition of the wave maps nonlinearity at fixed frequency 2^k ,

$$\begin{aligned} & P_k(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi) \\ (5.2) \quad & = 2 \sum_{k_2 \leq k-10} \phi_{\leq k_2-10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_k \\ (5.3) \quad & + \sum_{k_2 > k-10} \sum_{k_3 = k_2 + O(1)} P_k(\phi_{\leq k_2+20} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_{k_3}) \\ (5.4) \quad & + 2 \sum_{k_2 \leq k-10} \phi_{\leq k_2-10} (P_k(\partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_{k-5 < \cdot < k+5}) - \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_k). \\ (5.5) \quad & + \sum_{k_2 \leq k-10} \sum_{k_3 = k + O(1)} P_k L(\phi_{k_2-10 < \cdot \leq k-10}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \\ (5.6) \quad & + \sum_{k_2 > k-10} \sum_{k_1 = k_2 + O(1)} \sum_{k_3 \leq k_2 + O(1)} P_k L(\phi_{k_1}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \\ (5.7) \quad & + \sum_{k_2 \leq k-10} \sum_{k_3 = k + O(1)} 2^{-k} P_k L(\nabla_x \phi_{\leq k_2-10}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}). \end{aligned}$$

In order to arrive at this decomposition, we begin with

$$P_k(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi) = P_k(\phi \partial_\alpha \phi^\dagger_{\leq k-10} \partial^\alpha \phi_{\leq k-10}) + 2P_k(\phi \partial_\alpha \phi^\dagger_{\leq k-10} \partial^\alpha \phi_{> k-10}) + P_k(\phi \partial_\alpha \phi^\dagger_{> k-10} \partial^\alpha \phi_{> k-10})$$

and then further decompose into

$$\begin{aligned}
(5.8) \quad & P_k(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi) \\
(5.9) \quad & = P_k(\phi \partial_\alpha \phi^\dagger_{\leq k-10} \partial^\alpha \phi_{\leq k-10}) \\
(5.10) \quad & + 2 \sum_{k_1 > k-10} P_k(\phi_{k_1} \partial_\alpha \phi^\dagger_{\leq k-10} \partial^\alpha \phi_{k_1+O(1)}) \\
(5.11) \quad & + 2 \sum_{k_2 \leq k-10} P_k(\phi_{k_2-10 < \cdot \leq k-10} \partial_\alpha \phi^\dagger_{k_2} \partial^\alpha \phi_{k-5 < \cdot < k+5}) \\
(5.12) \quad & + 2 \sum_{k_2 \leq k-10} P_k(\phi_{\leq k_2-10} \partial_\alpha \phi^\dagger_{k_2} \partial^\alpha \phi_{k-5 < \cdot < k+5}) \\
(5.13) \quad & + \sum_{k_2 > k-10} P_k(\phi_{k_2-5 < \cdot < k_2+5} \partial_\alpha \phi^\dagger_{k_2} \partial^\alpha \phi_{\leq k_2-10}) \\
(5.14) \quad & + \sum_{k_2 > k-10} \sum_{k_3 > k_2+10} P_k(\phi_{k_3-5 < \cdot < k_3+5} \partial_\alpha \phi^\dagger_{k_2} \partial^\alpha \phi_{k_3}).
\end{aligned}$$

The first term (5.8) can be estimated in the same manner as the term (5.6) and is therefore not further included in the decomposition (5.2)–(5.7). The second term (5.9) is of type (5.6), while the third term (5.10) is of type (5.5). Using the Leibniz rule (2.1) for the projection P_k , we may write the fourth term (5.11) as

$$\begin{aligned}
(5.15) \quad & 2 \sum_{k_2 \leq k-10} P_k(\phi_{k_2-10 < \cdot \leq k-10} \partial_\alpha \phi^\dagger_{k_2} \partial^\alpha \phi_{k-5 < \cdot < k+5}) \\
& = 2 \sum_{k_2 \leq k-10} \phi_{\leq k_2-10} \partial_\alpha \phi^\dagger_{k_2} \partial^\alpha \phi_k \\
& + 2 \sum_{k_2 \leq k-10} 2^{-k} P_k L(\nabla_x \phi_{\leq k_2-10}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k-5 < \cdot < k+5}) \\
& + 2 \sum_{k_2 \leq k-10} \phi_{\leq k_2-10} (P_k(\partial_\alpha \phi^\dagger_{k_2} \partial^\alpha \phi_{k-5 < \cdot < k+5}) - \partial_\alpha \phi^\dagger_{k_2} \partial^\alpha \phi_k).
\end{aligned}$$

Then the first term on the right-hand side of (5.15) coincides exactly with the term (5.2), the second term on the right-hand side is of type (5.7) and the third term on the right-hand side coincides with the term (5.4). Finally, the terms (5.12) and (5.14) are both of the type (5.6), while the high-high interactions term (5.13) coincides with the term (5.3).

Let us now return to the decomposition (5.2)–(5.7) of the wave maps nonlinearity at fixed frequency. In order to estimate the term (5.5) we will use the trilinear estimate (4.4), while we will bound the terms (5.6) and (5.7) using a combination of the null form estimate (4.1) and Strichartz estimates. In contrast, the first three terms (5.2) – (5.4) are not amenable to good bounds using the null form estimate and the trilinear estimate. Correspondingly, we have to introduce a renormalization procedure to deal with these non-perturbative terms. For the term (5.2) we will follow Tao’s idea [36, 37] to apply a suitable gauge transformation that casts this part of the nonlinearity into a better form, while for the other two terms (5.3)–(5.4) we will pass to a “nonlinearly modified” version of ϕ_k that satisfies a better equation.

We begin with the latter part of our renormalization procedure. Using the identity

$$(5.16) \quad \begin{aligned} \square(\phi^{(1)}\phi^{(2)\dagger}\phi^{(3)}) &= 2\phi^{(1)}\partial_\alpha\phi^{(2)\dagger}\partial^\alpha\phi^{(3)} + 2\partial_\alpha\phi^{(1)}\partial^\alpha\phi^{(2)\dagger}\phi^{(3)} + 2\partial_\alpha\phi^{(1)}\phi^{(2)\dagger}\partial^\alpha\phi^{(3)} \\ &\quad + \square\phi^{(1)}\phi^{(2)\dagger}\phi^{(3)} + \phi^{(1)}\square\phi^{(2)\dagger}\phi^{(3)} + \phi^{(1)}\phi^{(2)\dagger}\square\phi^{(3)}, \end{aligned}$$

we may rewrite the high-high interaction term (5.3) into the schematic form

$$(5.17) \quad \sum_{k_2 > k-10} \sum_{k_3 = k_2 + O(1)} P_k(\phi_{\leq k_2+20} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_{k_3}) = \square \left(\frac{1}{2} \sum_{k_2 > k-10} \sum_{k_3 = k_2 + O(1)} P_k(\phi_{\leq k_2+20} \phi_{k_2}^\dagger \phi_{k_3}) \right) \\ + \sum_{k_2 > k-10} \sum_{k_3 = k_2 + O(1)} P_k L(\partial_\alpha \phi_{\leq k_2+20}, \partial^\alpha \phi_{k_2}, \phi_{k_3})$$

$$(5.18) \quad + \sum_{k_2 > k-10} \sum_{k_3 = k_2 + O(1)} P_k L(\square \phi_{\leq k_2+20}, \phi_{k_2}, \phi_{k_3})$$

$$(5.19) \quad + \sum_{k_2 > k-10} \sum_{k_3 = k_2 + O(1)} P_k L(\phi_{\leq k_2+20}, \square \phi_{k_2}, \phi_{k_3}).$$

Here we note that all terms apart from the first one on the right-hand side of the previous equation now have a better structure in that a derivative falls on a lowest frequency or upon reinserting the wave maps equation these terms become quintilinear and thus easier to estimate. In particular, we note that the term (5.17) is of the same form as the term (5.6).

Similarly, using the Leibniz rule (2.1) for the projection P_k and the identity (5.16), the other difficult term (5.4) can be rewritten into the schematic form

$$(5.20) \quad 2 \sum_{k_2 \leq k-10} \phi_{\leq k_2-10} (P_k(\partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_{k-5 < \cdot < k+5}) - \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_k) = \square \left(\sum_{k_2 \leq k-10} \phi_{\leq k_2-10} (P_k(\phi_{k_2}^\dagger \phi_{k-5 < \cdot < k+5}) - \phi_{k_2}^\dagger \phi_k) \right) \\ + \sum_{k_2 \leq k-10} 2^{-k} L(\partial_\alpha \phi_{\leq k_2-10}, \nabla_x \partial^\alpha \phi_{k_2}, \phi_{k-5 < \cdot < k+5})$$

$$(5.21) \quad + \sum_{k_2 \leq k-10} 2^{-k} L(\partial_\alpha \phi_{\leq k_2-10}, \nabla_x \phi_{k_2}, \partial^\alpha \phi_{k-5 < \cdot < k+5})$$

$$(5.22) \quad + \sum_{k_2 \leq k-10} 2^{-k} L(\square \phi_{\leq k_2-10}, \nabla_x \phi_{k_2}, \phi_{k-5 < \cdot < k+5})$$

$$(5.23) \quad + \sum_{k_2 \leq k-10} 2^{-k} L(\phi_{\leq k_2-10}, \nabla_x \square \phi_{k_2}, \phi_{k-5 < \cdot < k+5})$$

$$(5.24) \quad + \sum_{k_2 \leq k-10} 2^{-k} L(\phi_{\leq k_2-10}, \nabla_x \phi_{k_2}, \square \phi_{k-5 < \cdot < k+5}),$$

where again all terms apart from the first one on the right-hand side are now amenable to better estimates either because a derivative now falls on the lowest frequency or because upon reinserting the wave maps equation the term is now quintilinear. We observe that the terms (5.20) and (5.21) are essentially of the same form as the term (5.7) and can be dealt with in the same manner. Hence, introducing the following “nonlinearly modified” version of ϕ_k ,

$$(5.25) \quad \bar{\phi}_k := \phi_k + \frac{1}{2} \sum_{k_2 > k-10} \sum_{k_3 = k_2 + O(1)} P_k(\phi_{\leq k_2+20} \phi_{k_2}^\dagger \phi_{k_3}) + \sum_{k_2 \leq k-10} \phi_{\leq k_2-10} (P_k(\phi_{k_2}^\dagger \phi_{k-5 < \cdot < k+5}) - \phi_{k_2}^\dagger \phi_k),$$

our previous considerations imply that $\bar{\phi}_k$ satisfies a wave equation of the following schematic form

$$\begin{aligned}
(5.26) \quad \square \bar{\phi}_k = & -2 \sum_{k_2 \leq k-10} \phi_{\leq k_2-10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_k \\
& + \sum_{k_2 \leq k-10} \sum_{k_3=k+O(1)} P_k L(\phi_{k_2-10 < \cdot \leq k-10}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \\
& + \sum_{k_2 > k-10} \sum_{k_1=k_2+O(1)} \sum_{k_3 \leq k_2+O(1)} P_k L(\phi_{k_1}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \\
& + \sum_{k_2 \leq k-10} \sum_{k_3=k+O(1)} 2^{-k} P_k L(\nabla_x \phi_{\leq k_2-10}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \\
& + Q_k(\phi),
\end{aligned}$$

where $Q_k(\phi)$ comprises all the quintilinear terms (5.18)–(5.19) and (5.22)–(5.24)

$$\begin{aligned}
Q_k(\phi) = & \sum_{k_2 > k-10} \sum_{k_3=k_2+O(1)} P_k L(P_{\leq k_2+20}(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi), \phi_{k_2}, \phi_{k_3}) \\
& + \sum_{k_2 > k-10} \sum_{k_3=k_2+O(1)} P_k L(\phi_{\leq k_2+20}, P_{k_2}(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi), \phi_{k_3}) \\
& + \sum_{k_2 \leq k-10} 2^{-k} L(P_{\leq k_2-10}(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi), \nabla_x \phi_{k_2}, \phi_{k-5 < \cdot < k+5}) \\
& + \sum_{k_2 \leq k-10} 2^{-k} L(\phi_{\leq k_2-10}, \nabla_x P_{k_2}(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi), \phi_{k-5 < \cdot < k+5}) \\
& + \sum_{k_2 \leq k-10} 2^{-k} L(\phi_{\leq k_2-10}, \nabla_x \phi_{k_2}, P_{k-5 < \cdot < k+5}(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi)).
\end{aligned}$$

All the quintilinear terms in $Q_k(\phi)$ can be easily estimated using just combinations of the null form estimate (4.1) and Strichartz estimates as well as the fact that $\|\phi\|_{L_t^\infty L_x^\infty} \lesssim 1$. Thus, we can overcome the difficulties with the terms (5.3)–(5.4) at the expense of passing from ϕ_k to the new variable $\bar{\phi}_k$. Note that $\bar{\phi}_k$ is also localized to frequency 2^k . In the following we will sometimes use the notation $\bar{\phi} = \sum_{k \in \mathbb{Z}} \bar{\phi}_k$.

Now we still have to deal with the more severe term (5.2) which comes up as the first term on the right-hand side of the equation (5.26) for $\bar{\phi}_k$. We follow Tao's idea [36, 37] and apply a renormalization $U_{\leq k-10} \bar{\phi}_k$ so that $\square(U_{\leq k-10} \bar{\phi}_k)$ takes on a much better form. More specifically, we implement a version of the gauge construction introduced by Sterbenz-Tataru [29, 30]. To this end we pick a smooth cut-off function $\eta \in C^\infty(\mathbb{R})$ such that $\eta(y) = 0$ for $y \leq -\frac{2}{3}$ and $\eta(y) = 1$ for $y \geq -\frac{1}{3}$. Then we define for $h \in \mathbb{R}$ the matrix

$$(5.27) \quad B_h := \frac{d}{dh} \sum_{\ell \in \mathbb{Z}} \eta(h - \ell) (\phi_{\leq \ell-10} \phi_\ell^\dagger - \phi_\ell \phi_{\leq \ell-10}^\dagger) = \sum_{\ell \in \mathbb{Z}} \eta'(h - \ell) (\phi_{\leq \ell-10} \phi_\ell^\dagger - \phi_\ell \phi_{\leq \ell-10}^\dagger).$$

Now we define the gauge transformations $U_{\leq h}$ for $h \in \mathbb{R}$ via the ODE

$$(5.28) \quad \frac{d}{dh} U_{\leq h} = U_{\leq h} B_h, \quad \lim_{h \rightarrow -\infty} U_{\leq h} = Id.$$

By the anti-symmetry of B_h , it follows that the gauge transformations $U_{\leq h}$ as solutions to the ODE (5.28) satisfy $U_{\leq h}^\dagger U_{\leq h} = Id$ and are thus exactly orthogonal. However, they do not have a sharp frequency localization. Fortunately, we will see that the renormalized quantity $U_{\leq k-10} \bar{\phi}_k$ is still approximately localized to frequency 2^k up to exponentially decaying tails.

Next, we compute the equation that the renormalization $U_{\leq k-10} \bar{\phi}_k$ satisfies. Upon defining for $k \in \mathbb{Z}$

$$(5.29) \quad A_{\alpha;k} = \partial_\alpha \phi_k \phi_{\leq k-10}^\dagger - \phi_{\leq k-10} \partial_\alpha \phi_k^\dagger$$

and

$$(5.30) \quad A_{\alpha;\leq k} = \sum_{\ell \leq k} A_{\alpha;\ell},$$

we find that

$$(5.31) \quad \begin{aligned} \square(U_{\leq k-10}\bar{\phi}_k) &= U_{\leq k-10}(\square\bar{\phi}_k - 2A_{\alpha;\leq k-10}\partial^\alpha\phi_k) \\ &\quad + 2(\partial_\alpha U_{\leq k-10} + U_{\leq k-10}A_{\alpha;\leq k-10})\partial^\alpha\phi_k \\ &\quad + 2\partial_\alpha U_{\leq k-10}\partial^\alpha(\bar{\phi}_k - \phi_k) \\ &\quad + (\square U_{\leq k-10})\bar{\phi}_k. \end{aligned}$$

In the following proposition we carefully analyze the structure of each term in the nonlinearity of the wave equation (5.31) for $U_{\leq k-10}\bar{\phi}_k$.

Proposition 5.1. *It holds that*

$$(5.32) \quad \begin{aligned} \square\bar{\phi}_k - 2A_{\alpha;\leq k-10}\partial^\alpha\phi_k &= \sum_{k_2 \leq k-10} \sum_{k_3=k+O(1)} P_k L(\phi_{k_2-10}, \partial_\alpha\phi_{k_2}, \partial^\alpha\phi_{k_3}) \\ &\quad + \sum_{k_2 > k-10} \sum_{k_1=k_2+O(1)} \sum_{k_3 \leq k_2+O(1)} P_k L(\phi_{k_1}, \partial_\alpha\phi_{k_2}, \partial^\alpha\phi_{k_3}) \\ &\quad + \sum_{k_2 \leq k-10} \sum_{k_3=k+O(1)} 2^{-k} P_k L(\nabla_x \phi_{\leq k_2-10}, \partial_\alpha\phi_{k_2}, \partial^\alpha\phi_{k_3}) \\ &\quad + Q_k(\phi). \end{aligned}$$

Moreover, we have the following schematic identities

$$(5.33) \quad (\partial_\alpha U_{\leq k-10} + U_{\leq k-10}A_{\alpha;\leq k-10})\partial^\alpha\phi_k = \sum_{k_2 \leq k-10} U_{\leq k_2-1} L(\phi_{k_2-10}, \partial_\alpha\phi_{k_2}, \partial^\alpha\phi_k),$$

$$(5.34) \quad \partial_\alpha U_{\leq k-10}\partial^\alpha(\bar{\phi}_k - \phi_k) = \sum_{k_1 \leq k-10} U_{\leq k_1-1} L(\partial_\alpha L(\phi_{\leq k_1-10}, \phi_{k_1}), \partial^\alpha(\bar{\phi}_k - \phi_k)),$$

$$(5.35) \quad \begin{aligned} (\square U_{\leq k-10})\bar{\phi}_k &= \sum_{k_1 \leq k-10} U_{\leq k_1-1} L(\square L(\phi_{\leq k_1-10}, \phi_{k_1}), \bar{\phi}_k) \\ &\quad + \sum_{k_1 \leq k_2 \leq k-10} U_{\leq k_1-1} L(\partial_\alpha L(\phi_{\leq k_1-10}, \phi_{k_1}), \partial^\alpha L(\phi_{\leq k_2-10}, \phi_{k_2}), \bar{\phi}_k). \end{aligned}$$

Proof. We begin with the proof of the first identity (5.32). In view of (5.26) we only have to understand the structure of the term

$$\sum_{k_2 \leq k-10} \partial_\alpha\phi_{k_2}\phi_{\leq k_2-10}^\dagger \partial^\alpha\phi_k.$$

But here we can argue as in Tao [37, p. 461] and use the geometric identity $\phi^\dagger \partial^\alpha\phi = 0$ to show that this term is in fact a sum of terms of the schematic forms (5.5)–(5.7). This proves (5.32).

Next we turn to the proof of the schematic identity (5.33). Using the exact orthogonality of the gauge transformations $U_{\leq h}$, by telescoping we may write $U_{\leq k-10}$ as

$$U_{\leq k-10} = U_{\leq -M} \prod_{j=k-M+1}^{k-10} U_{\leq j-1}^\dagger U_{\leq j}$$

for some $M \gg 1$. Applying a derivative ∂_α and letting $M \rightarrow \infty$, we find that

$$\partial_\alpha U_{\leq k-10} = \sum_{j \leq k-10} U_{\leq j-1} \partial_\alpha (U_{\leq j-1}^\dagger U_{\leq j}) U_{\leq j}^\dagger U_{\leq k-10}.$$

From the defining ODE (5.28) for the gauge transformations we have that

$$(5.36) \quad U_{\leq j-1}^\dagger U_{\leq j} = \int_{j-1}^j U_{\leq j-1}^\dagger U_{\leq h} B_h dh + Id$$

and hence

$$\partial_\alpha (U_{\leq j-1}^\dagger U_{\leq j}) = \partial_\alpha \int_{j-1}^j B_h dh + \partial_\alpha \int_{j-1}^j (U_{\leq j-1}^\dagger U_{\leq j} - Id) B_h dh.$$

Thus, we may write

$$\begin{aligned} \partial_\alpha U_{\leq k-10} + U_{\leq k-10} A_{\alpha; \leq k-10} &= \sum_{j \leq k-10} U_{\leq j-1} \left(\partial_\alpha \int_{j-1}^j B_h dh \right) U_{\leq j}^\dagger U_{\leq k-10} \\ &+ \sum_{j \leq k-10} U_{\leq j-1} A_{\alpha; j} \\ &+ \sum_{j \leq k-10} U_{\leq j-1} (U_{\leq j-1}^\dagger U_{\leq k-10} - Id) A_{\alpha; j} \\ &+ \sum_{j \leq k-10} U_{\leq j-1} \left(\partial_\alpha \int_{j-1}^j (U_{\leq j-1}^\dagger U_{\leq h} - Id) B_h dh \right) U_{\leq j}^\dagger U_{\leq k-10}. \end{aligned}$$

Now by the definition of B_h (5.27) we have that

$$\partial_\alpha \int_{j-1}^j B_h dh = \partial_\alpha (\phi_{\leq j-10} \phi_j^\dagger - \phi_j \phi_{\leq j-10}^\dagger) = -A_{\alpha; j} + (\partial_\alpha \phi_{\leq j-10} \phi_j^\dagger - \phi_j \partial_\alpha \phi_{\leq j-10}^\dagger).$$

Combining the two previous identities we find that

$$\begin{aligned} \partial_\alpha U_{\leq k-10} + U_{\leq k-10} A_{\alpha; \leq k-10} &= \sum_{j \leq k-10} U_{\leq j-1} A_{\alpha; j} (-U_{\leq j}^\dagger U_{\leq k-10} + Id) \\ &+ \sum_{j \leq k-10} U_{\leq j-1} (\partial_\alpha \phi_{\leq j-10} \phi_j^\dagger - \phi_j \partial_\alpha \phi_{\leq j-10}^\dagger) U_{\leq j}^\dagger U_{\leq k-10} \\ (5.37) \quad &+ \sum_{j \leq k-10} U_{\leq j-1} (U_{\leq j-1}^\dagger U_{\leq k-10} - Id) A_{\alpha; j} \\ &+ \sum_{j \leq k-10} U_{\leq j-1} \left(\partial_\alpha \int_{k-1}^k (U_{\leq j-1}^\dagger U_{\leq h} - Id) B_h dh \right) U_{\leq j}^\dagger U_{\leq k-10}. \end{aligned}$$

In view of the identity

$$U_{\leq k_1}^\dagger U_{\leq k_2} - Id = \int_{k_1}^{k_2} U_{\leq k_1}^\dagger U_{\leq h} B_h dh \quad \text{for } k_1 < k_2,$$

the exact orthogonality of the gauge transformations $U_{\leq h}$, the definition of B_h (5.27) and the fact that $\|\phi\|_{L_t^\infty L_x^\infty} \lesssim 1$, it is now apparent that when it comes to estimates, the first three terms on the right-hand side of (5.37) are of the schematic form (5.33). To see this also for the last term on the right-hand side of (5.37), we note that by iteratively inserting the identity (5.36), we obtain that

$$\begin{aligned} \partial_\alpha \int_{k-1}^k (U_{\leq j-1}^\dagger U_{\leq h} - Id) B_h dh &= \int_{k-1}^k \int_{k-1}^{h_1} U_{\leq j-1}^\dagger U_{\leq h_2} B_{h_2} \partial_\alpha B_{h_1} dh_2 dh_1 \\ &+ \sum_{n=1}^\infty \int_{k-1}^k \int_{k-1}^{h_1} \cdots \int_{k-1}^{h_n} U_{\leq j-1}^\dagger U_{\leq h_{n+1}} \partial_\alpha B_{h_{n+1}} B_{h_n} \cdots B_{h_1} dh_{n+1} \cdots dh_1. \end{aligned}$$

Here the series may seem to diverge at first sight, but the integration over simplices yields the necessary compensating $\frac{1}{n!}$ decay. Thus, also the last term on the right-hand side of (5.37) is of the schematic form (5.33), which finishes the proof of (5.33).

Similarly, we can prove (5.34) using the identities

$$\partial_\alpha U_{\leq k-10} = \sum_{j \leq k-10} U_{\leq j-1} \partial_\alpha (U_{\leq j-1}^\dagger U_{\leq j}) U_{\leq j}^\dagger U_{\leq k-10}$$

and

$$\begin{aligned} \partial_\alpha (U_{\leq j-1}^\dagger U_{\leq j}) &= \int_{j-1}^j U_{\leq j-1}^\dagger U_{\leq h_1} \partial_\alpha B_{h_1} dh_1 \\ &+ \sum_{n=1}^{\infty} \int_{j-1}^j \int_{j-1}^{h_1} \cdots \int_{j-1}^{h_n} U_{\leq j-1}^\dagger U_{\leq h_{n+1}} \partial_\alpha B_{h_{n+1}} B_{h_n} \cdots B_{h_1} dh_{n+1} \cdots dh_1. \end{aligned}$$

Finally, we turn to the proof of (5.35). From the telescoping identity

$$U_{\leq k-10} = U_{\leq -M} \prod_{j=k-M+1}^{k-10} U_{\leq j-1}^\dagger U_{\leq j}$$

for $M \gg 1$, we obtain that

$$\begin{aligned} (5.38) \quad \square U_{\leq k-10} &= \sum_{j \leq k-10} U_{\leq j-1} \square (U_{\leq j-1}^\dagger U_{\leq j}) U_{\leq j}^\dagger U_{\leq k-10} \\ &+ 2 \sum_{j_1 < j_2 \leq k-10} U_{\leq j_1-1} \partial_\alpha (U_{\leq j_1-1}^\dagger U_{\leq j_1}) \cdots \partial^\alpha (U_{\leq j_2-1}^\dagger U_{\leq j_2}) U_{\leq j_2}^\dagger U_{\leq k-10}. \end{aligned}$$

Then by iteratively inserting the identities

$$\square (U_{\leq j-1}^\dagger U_{\leq j}) = \int_{j-1}^j \square (U_{\leq j-1}^\dagger U_{\leq h}) B_h dh + 2 \int_{j-1}^j \partial_\alpha (U_{\leq j-1}^\dagger U_{\leq h}) \partial^\alpha B_h dh + \int_{j-1}^j (U_{\leq j-1}^\dagger U_{\leq h}) \square B_h dh$$

and

$$\partial_\alpha (U_{\leq j-1}^\dagger U_{\leq j}) = \int_{j-1}^j \partial_\alpha (U_{\leq j-1}^\dagger U_{\leq h}) B_h dh + \int_{j-1}^j (U_{\leq j-1}^\dagger U_{\leq h}) \partial_\alpha B_h dh$$

into (5.38), we infer by similar arguments as before that $(\square U_{\leq k-10}) \bar{\phi}_k$ is indeed of the schematic form (5.35). \square

Next, we prove several useful bounds on $\bar{\phi}_k$ and on the gauge transformations that will be needed in the sequel.

Lemma 5.2. *Let $\phi: I \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$ be a radial wave map of energy E . Then we have for all $k \in \mathbb{Z}$ that*

$$(5.39) \quad \|\bar{\phi}_k\|_{L_t^\infty L_x^\infty} \lesssim 1 + E$$

and it holds that

$$(5.40) \quad \sum_{k \in \mathbb{Z}} \|\nabla_{t,x} \bar{\phi}_k\|_{L_x^2}^2 \lesssim E + E^3.$$

Proof. We start with the proof of (5.39). Since ϕ maps into the sphere and the Littlewood-Paley projections P_k are disposable, we clearly have $\|\phi_k\|_{L_t^\infty L_x^\infty} \lesssim 1$. Using Bernstein's inequality, we easily bound the second component in the definition (5.25) of $\tilde{\phi}_k$ by

$$\begin{aligned} \left\| \sum_{k_2 > k-10} \sum_{k_3 = k_2 + O(1)} P_k(\phi_{\leq k_2+20} \phi_{k_2}^\dagger \phi_{k_3}) \right\|_{L_x^\infty} &\lesssim \sum_{k_2 > k-10} \sum_{k_3 = k_2 + O(1)} 2^{2k} \|\phi_{k_2}\|_{L_x^2} \|\phi_{k_3}\|_{L_x^2} \\ &\lesssim \sum_{k_2 > k-10} \|\nabla_x \phi_{k_2}\|_{L_x^2}^2 \\ &\lesssim E, \end{aligned}$$

while for the third component in the definition (5.25) we additionally use the Leibniz rule (2.1) for P_k to find that

$$\begin{aligned} \left\| \sum_{k_2 \leq k-10} \phi_{\leq k_2-10} (P_k(\phi_{k_2}^\dagger \phi_{k-5 < \cdot < k+5}) - \phi_{k_2}^\dagger \phi_k) \right\|_{L_x^\infty} &= \left\| \sum_{k_2 \leq k-10} 2^{-k} L(\phi_{\leq k_2-10}, \nabla_x \phi_{k_2}, \phi_{k+O(1)}) \right\|_{L_x^\infty} \\ &\lesssim \sum_{k_2 \leq k-10} 2^{-k+k_2} \|\nabla_x \phi_{k_2}\|_{L_x^2} \|\nabla_x \phi_k\|_{L_x^2} \\ &\lesssim E. \end{aligned}$$

The proof of (5.40) follows by analogous estimates. \square

Lemma 5.3. *Let $\phi: I \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$ be a radial wave map of energy E and let $U_{\leq h}$ for $h \in \mathbb{R}$ be the associated gauge transformations as defined in (5.28). For $2 < p \leq \infty$ it holds that*

$$(5.41) \quad \|\nabla_{t,x} U_{\leq k}\|_{L_t^\infty L_x^p} \lesssim 2^{(1-\frac{2}{p})k} (E^{\frac{1}{2}} + E).$$

Moreover, we have for $2 \leq p \leq \infty$ that

$$(5.42) \quad \|\nabla_x \nabla_{t,x} U_{\leq k}\|_{L_t^\infty L_x^p} \lesssim 2^{(2-\frac{2}{p})k} (E^{\frac{1}{2}} + E).$$

Proof. By telescoping as in the proof of Proposition 5.1, we obtain the schematic identity

$$\nabla_{t,x} U_{\leq k} = \sum_{k_2 \leq k} U_{\leq k_2-1} \nabla_{t,x} L(\phi_{\leq k_2-10}, \phi_{k_2}) U_{\leq k_2}^\dagger U_{\leq k}.$$

Hence we find for $2 < p \leq \infty$ that

$$\begin{aligned} \|\nabla_{t,x} U_{\leq k}\|_{L_x^p} &\lesssim \sum_{k_2 \leq k} \|L(\nabla_{t,x} \phi_{\leq k_2-10}, \phi_{k_2})\|_{L_x^p} + \sum_{k_2 \leq k} \|L(\phi_{\leq k_2-10}, \nabla_{t,x} \phi_{k_2})\|_{L_x^p} \\ &\lesssim \sum_{k_2 \leq k} \sum_{k_1 \leq k_2-10} \|\nabla_{t,x} \phi_{k_1}\|_{L_x^\infty} \|\phi_{k_2}\|_{L_x^p} + \sum_{k_2 \leq k} \|\nabla_{t,x} \phi_{k_2}\|_{L_x^p} \\ &\lesssim \sum_{k_2 \leq k} \sum_{k_1 \leq k_2-10} 2^{k_1} 2^{-\frac{2}{p}k_2} \|\nabla_{t,x} \phi_{k_1}\|_{L_x^2} \|\nabla_x \phi_{k_2}\|_{L_x^2} + \sum_{k_2 \leq k} 2^{(1-\frac{2}{p})k_2} \|\nabla_{t,x} \phi_{k_2}\|_{L_x^2} \\ &\lesssim 2^{(1-\frac{2}{p})k} (E + E^{\frac{1}{2}}), \end{aligned}$$

which yields (5.41). The proof of the bound (5.42) proceeds similarly. \square

We conclude this section by establishing a small energy global regularity result for the wave maps equation (WM) for radially symmetric, classical initial data. The proof is a fairly immediate consequence of the multilinear estimates from the previous section and of the careful decomposition of the wave maps nonlinearity as well as of the introduction of the renormalization procedure in this section.

Theorem 5.4. *There exists an absolute constant $\varepsilon > 0$ such that for any radially symmetric, classical initial data $(\phi_0, \phi_1): \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ with energy $E[\phi] < \varepsilon$, there exists a unique, radially symmetric, classical global solution $\phi: \mathbb{R}^{1+2} \rightarrow \mathbb{S}^m$ to (WM) with initial data $\phi[0] = (\phi_0, \phi_1)$ satisfying*

$$(5.43) \quad \|\phi\|_{S[\mathbb{R}]} \lesssim E[\phi]^{\frac{1}{2}}.$$

Proof. It is a standard argument to show that in order to conclude global regularity and the bound (5.43), it suffices to prove frequency envelope bounds for the local-in-time evolution of a classical radial wave map with initial data (ϕ_0, ϕ_1) , see [36, 37]. Thus, let $\{c_k\}_{k \in \mathbb{Z}}$ be a frequency envelope covering the initial data (ϕ_0, ϕ_1) and denote by ϕ be the classical wave maps evolution with initial data $\phi[0] = (\phi_0, \phi_1)$ on a time interval $I = [-T, T]$ for some $0 < T < \infty$. We shall prove frequency envelope bounds for the evolution by a bootstrap argument. Making the bootstrap assumption

$$\|P_k \phi\|_{S_k[I]} \leq C c_k, \quad k \in \mathbb{Z},$$

for some sufficiently large, absolute constant $C > 0$, we shall now show that this implies the improved bound

$$\|P_k \phi\|_{S_k[I]} \leq \frac{C}{2} c_k, \quad k \in \mathbb{Z}.$$

By scaling invariance, it suffices to prove this bound for $k = 0$. To this end we first deduce a frequency envelope bound for the renormalized quantity $U_{\leq -10} \bar{\phi}_0$. This step is slightly compounded by the fact that the variable $U_{\leq -10} \bar{\phi}_0$ is only approximately localized to frequency ~ 0 up to exponentially decaying tails. Correspondingly, we show that for any $0 \leq \alpha < 1$, it holds that

$$(5.44) \quad \sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k_0|} \|P_{k_0}(U_{\leq -10} \bar{\phi}_0)\|_{S_{k_0}[I]} \ll C c_0.$$

From this bound on $U_{\leq -10} \bar{\phi}_0$ we can then pass back to ϕ_0 and recover the improved bound

$$\|P_0 \phi\|_{S_0[I]} \leq \frac{C}{2} c_0,$$

as explained in detail in the proof of Proposition 6.3 in the next section. In order to prove the bound (5.44) we show that for any $0 \leq \alpha < 1$,

$$(5.45) \quad \sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k_0|} \|P_{k_0} \nabla_{t,x}(U_{\leq -10} \bar{\phi}_0)\|_{L_x^2} \ll C c_0$$

and that

$$(5.46) \quad \sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k_0|} \|P_{k_0} \square(U_{\leq -10} \bar{\phi}_0)\|_{L_t^1 L_x^2[I]} \lesssim \varepsilon C c_0.$$

For sufficiently small $\varepsilon > 0$ the energy estimate (3.1) then yields the desired bound (5.44). We note that the exponential factors $2^{\alpha|k_0|}$ in (5.45) and (5.46) can be controlled by playing out Bernstein's inequality and the bounds on the gauge transformations from Lemma 5.3, see the proof of Proposition 6.3 in the next section for such an argument. Then the bound (5.45) is straightforward to derive from the definition of $\bar{\phi}_0$ using the properties of frequency envelopes. Finally, in order to deduce the bound (5.46), we recall that Proposition 5.1 carefully lists each schematic term arising in the nonlinearity $\square(U_{\leq -10} \bar{\phi}_0)$. Using the properties of frequency envelopes we then prove (5.46) separately for each schematic term. Specifically, we may bound the first term on the right-hand side of (5.32) as well as the term (5.33) using the trilinear estimate (4.4), while all other terms on the right-hand side of (5.32) as well as the terms (5.34)–(5.35) can be estimated by a combination of the null form estimate (4.1) and Strichartz estimates. \square

6. BREAKDOWN CRITERION

Here we show that the S norm introduced in Definition 3.1 provides sufficient control on a radially symmetric, classical solution to the wave maps equation (WM) in order to infer long time existence and scattering. The main result is summarized in the following proposition.

Proposition 6.1. *Let I be the maximal time interval of existence of a radially symmetric, classical wave map $\phi: I \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$. If $\|\phi\|_{S[I]} < \infty$, then we must have $I = \mathbb{R}$. Moreover, ϕ then scatters to finite energy free waves as $t \rightarrow \pm\infty$ in the sense that there exist $(f_{\pm}, g_{\pm}) \in \dot{H}_x^1 \times L_x^2$ such that*

$$\lim_{t \rightarrow \pm\infty} \|\nabla_{t,x}\phi - \nabla_{t,x}S(t)(f_{\pm}, g_{\pm})\|_{L_x^2} = 0,$$

where $S(t)(f_{\pm}, g_{\pm}) = \cos(t|\nabla|)f_{\pm} + \frac{\sin(t|\nabla|)}{|\nabla|}g_{\pm}$ denotes the free wave propagator.

We will give the proof of Proposition 6.1 at the end of this section. The key ingredient for the proof is the next proposition which provides frequency envelope bounds for a radially symmetric, classical wave map $\phi: I \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$ with finite S norm.

Proposition 6.2. *Let $\phi: I \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$ be a radially symmetric, classical wave map defined on a time interval I containing time $t = 0$. Let $\{c_k\}_{k \in \mathbb{Z}}$ be a frequency envelope covering the initial data $\phi[0]$. If $\|\phi\|_{S[I]} < \infty$, then there exists $C \equiv C(\|\phi\|_{S[I]}) > 0$ such that for all $k \in \mathbb{Z}$,*

$$\|P_k\phi\|_{S_k[I]} \leq Cc_k.$$

The proof of Proposition 6.2 relies on the following weak divisibility property of the S norm.

Proposition 6.3. *Let $\phi: I \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$ be a radially symmetric, classical wave map of energy E defined on a time interval I and satisfying $\|\phi\|_{S[I]} \leq K$. Then there exists a partition $I = \cup_{j=1}^N I_j$ into $N \equiv N(K, E)$ consecutive time intervals I_j such that*

$$(6.1) \quad \|\phi\|_{S[I_j]} \lesssim C(E) \quad \text{for } j = 1, \dots, N,$$

where $C(E)$ is an absolute constant that just depends on the size of the energy E .

Proof. The proof proceeds in two steps. First we show that

$$(6.2) \quad \|\bar{\phi} - \phi\|_{S[I]} \lesssim C(E).$$

Then we prove that the interval I can be partitioned into $N(K, E)$ consecutive subintervals I_j such that

$$(6.3) \quad \|\bar{\phi}\|_{S[I_j]} \lesssim C(E) \quad \text{for } j = 1, \dots, N.$$

The assertion (6.1) then is an immediate consequence of the two previous bounds. Their proof, however, requires the introduction of certain frequency and spatial cutoffs in the definitions of the quantity $\bar{\phi}$ and of the gauge transformations $U_{\leq h}$ which we will explain as we go. We will be able to prove (6.2) directly, gaining smallness by choosing the frequency and spatial cutoff parameter sufficiently large, while we will establish (6.3) using a suitable divisibility argument.

We begin with the proof of the bound (6.2). To this end we recall that our motivation for the introduction of the quantity $\bar{\phi} = \sum_{k \in \mathbb{Z}} \bar{\phi}_k$ was in the first place that the high-high interaction term (5.3)

$$\sum_{k_2 > k-10} \sum_{k_3 = k_2 + O(1)} P_k(\phi_{\leq k_2+20} \partial_{\alpha} \phi_{k_2}^{\dagger} \partial^{\alpha} \phi_{k_3})$$

as well as the term (5.4)

$$\sum_{k_2 \leq k-10} \phi_{\leq k_2-10} (P_k(\partial_{\alpha} \phi_{k_2}^{\dagger} \partial^{\alpha} \phi_{k-5 < \cdot < k+5}) - \partial_{\alpha} \phi_{k_2}^{\dagger} \partial^{\alpha} \phi_k)$$

in the wave map nonlinearity localized to frequency $\sim 2^k$ turn out to be non-perturbative due to our choice to estimate the nonlinearity solely in $L_t^1 L_x^2$ and to build the S norm out of relatively elementary function

spaces. However, we can in fact “peel off” some more good parts from these terms. More precisely, the following frequency and spatially truncated versions of (5.3)

$$(6.4) \quad \sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} P_k(\chi_{\{r \leq 2^{-k+m}\}} \phi_{\leq k_2+20} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_{k_3})$$

and of (5.4)

$$(6.5) \quad 2 \sum_{k-m < k_2 \leq k-10} P_{k-10 < \cdot < k+10} (\chi_{\{r \leq 2^{-k+m}\}} \phi_{\leq k_2+20} (P_k(\partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_{k-5 < \cdot < k+5}) - \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_k))$$

for some large parameter $m \gg 1$, can still be suitably estimated, as we will show in detail in the second part of the proof of this proposition. Correspondingly, within this proof we may also work with the following modified definition

$$(6.6) \quad \bar{\phi}_k - \phi_k := \frac{1}{2} \sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} P_k(\chi_{\{r > 2^{-k+m}\}} \phi_{\leq k_2+20} \phi_{k_2}^\dagger \phi_{k_3})$$

$$(6.7) \quad + \frac{1}{2} \sum_{k_2 > k+m} \sum_{k_3=k_2+O(1)} P_k(\phi_{\leq k_2+20} \phi_{k_2}^\dagger \phi_{k_3})$$

$$(6.8) \quad + \sum_{k-m < k_2 \leq k-10} P_{k-10 < \cdot < k+10} (\chi_{\{r > 2^{-k+m}\}} \phi_{\leq k_2+20} (P_k(\phi_{k_2}^\dagger \phi_{k-5 < \cdot < k+5}) - \phi_{k_2}^\dagger \phi_k))$$

$$(6.9) \quad + \sum_{k_2 \leq k-m} P_{k-10 < \cdot < k+10} (\phi_{\leq k_2+20} (P_k(\phi_{k_2}^\dagger \phi_{k-5 < \cdot < k+5}) - \phi_{k_2}^\dagger \phi_k)).$$

Note that additional error terms will arise here in the equation for $\square \bar{\phi}_k$ when a derivative hits the spatial cutoffs $\chi_{\{r > 2^{-k+m}\}}$. However, these extra terms can be dealt with easily. We now prove that

$$(6.10) \quad \|\bar{\phi} - \phi\|_{S[I]} \lesssim 2^{-am} \|\phi\|_{S[I]}^\beta$$

for suitable $\alpha, \beta > 0$, which immediately implies the desired bound (6.2) upon choosing $m \gg 1$ sufficiently large depending on the size of $\|\phi\|_{S[I]}$. We start off with the term (6.6) in the expression for $\bar{\phi}_k - \phi_k$ and seek to show by direct estimation that

$$(6.11) \quad \left(\sum_{k \in \mathbb{Z}} \left\| \sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} P_k(\chi_{\{r > 2^{-k+m}\}} \phi_{\leq k_2+20} \phi_{k_2}^\dagger \phi_{k_3}) \right\|_{S_k[I]}^2 \right)^{\frac{1}{2}} \lesssim 2^{-am} \|\phi\|_{S[I]}^\beta.$$

To this end we prove (6.11) separately for each part of the S_k norm. We begin with the kinetic energy component. Suppose $\nabla_{t,x}$ hits $\phi_{k_2}^\dagger$. Using Strauss’ improved Sobolev embedding in the radial case

$$(6.12) \quad \|r^{\frac{1}{2}} \phi_k\|_{L_x^\infty} \lesssim 2^{-\frac{1}{2}k} \|\phi_k\|_{\dot{H}_x^1},$$

we may estimate

$$\begin{aligned} & \left(\sum_{k \in \mathbb{Z}} \left\| \sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} P_k(\chi_{\{r > 2^{-k+m}\}} \phi_{\leq k_2+20} \nabla_{t,x} \phi_{k_2}^\dagger \phi_{k_3}) \right\|_{L_t^\infty L_x^2[I]}^2 \right)^{\frac{1}{2}} \\ & \lesssim \left(\sum_{k \in \mathbb{Z}} \left(\sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} \|\nabla_{t,x} \phi_{k_2}\|_{L_t^\infty L_x^2[I]} \|\chi_{\{r > 2^{-k+m}\}} \phi_{k_3}\|_{L_t^\infty L_x^\infty[I]} \right)^2 \right)^{\frac{1}{2}} \\ & \lesssim \left(\sum_{k \in \mathbb{Z}} \left(\sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} \|\phi_{k_2}\|_{S_{k_2}[I]} 2^{-\frac{1}{2}m} 2^{-\frac{1}{2}(k_3-k)} \|\phi_{k_3}\|_{S_{k_3}[I]} \right)^2 \right)^{\frac{1}{2}} \\ & \lesssim 2^{-\frac{1}{2}m} \|\phi\|_{S[I]}^2. \end{aligned}$$

Analogously, we can deal with the expressions arising when $\nabla_{t,x}$ hits $\phi_{\leq k_2+20}$ or the cutoff $\chi_{\{r>2^{-k+m}\}}$, which yields (6.11) for the kinetic energy component of the S_k norm. In a very similar fashion we may also prove (6.11) for the Strichartz, weighted endpoint Strichartz and local energy decay components of the S_k norm. Hence we are left to prove (6.11) for the more delicate atomic space parts Z_k^\pm . It suffices to consider the Z_k^+ component. Assume that $(\partial_t + \partial_r)$ hits the input $\phi_{k_2}^\dagger$. Then we decompose $(\partial_t + \partial_r)\phi_{k_2}$ into its (+) and (-) components

$$(\partial_t + \partial_r)\phi_{k_2} = (\partial_t + \partial_r)\phi_{k_2}^{(+)} + (\partial_t + \partial_r)\phi_{k_2}^{(-)}.$$

We now show that in case of the $(\partial_t + \partial_r)\phi_{k_2}^{(+)}$ component, we can place the whole expression

$$\sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} P_k(\chi_{\{r>2^{-k+m}\}}\phi_{\leq k_2+20}(\partial_t + \partial_r)\phi_{k_2}^{(+)\dagger}\phi_{k_3})$$

into the (+) component of the Z_k^+ norm with the desired bound. Indeed, we have for the $L_{t+r}^2 L_{t-r}^\infty$ part that

$$\begin{aligned} & \left(\sum_{k \in \mathbb{Z}} \left(\sup_{0 < \lambda \leq \frac{1}{2}} 2^{-(\frac{1}{2}-\lambda)k} \left\| r^{+\lambda} \sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} P_k(\chi_{\{r>2^{-k+m}\}}\phi_{\leq k_2+20}(\partial_t + \partial_r)\phi_{k_2}^{(+)\dagger}\phi_{k_3}) \right\|_{L_{t+r}^2 L_{t-r}^\infty[I]} \right)^2 \right)^{\frac{1}{2}} \\ & \lesssim \left(\sum_{k \in \mathbb{Z}} \left(\sup_{0 < \lambda \leq \frac{1}{2}} \sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} 2^{-(\frac{1}{2}-\lambda)k} \left\| r^{+\lambda}(\partial_t + \partial_r)\phi_{k_2}^{(+)} \right\|_{L_{t+r}^2 L_{t-r}^\infty} \left\| \chi_{\{r>2^{-k+m}\}}\phi_{k_3} \right\|_{L_t^\infty L_x^\infty} \right)^2 \right)^{\frac{1}{2}} \\ & \lesssim \left(\sum_{k \in \mathbb{Z}} \left(\sup_{0 < \lambda \leq \frac{1}{2}} \sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} 2^{-\frac{1}{2}m} 2^{-\lambda(k_2-k)} \|\phi_{k_2}\|_{S_{k_2}} \|\phi_{k_3}\|_{S_{k_3}} \right)^2 \right)^{\frac{1}{2}} \\ & \lesssim 2^{-\frac{1}{2}m} m^{\frac{1}{2}} \|\phi\|_{S[I]}^2, \end{aligned}$$

where we again used Strauss' improved Sobolev embedding (6.12). Similarly, we may bound the Y_k norm.

Instead, when we face the $(\partial_t + \partial_r)\phi_{k_2}^{(-)}$ component, we can suitably place the whole resulting expression into the (-) component of the Z_k^+ norm. More precisely, for the weighted $L_t^q L_x^p$ norm we find that

$$\begin{aligned} & 2^{(\frac{1}{q}+\frac{2}{p}-1)k} 2^{\lambda k} \sum_{\ell \in \mathbb{Z}} \left\| \chi_{\{r \sim 2^\ell\}} r^{+\lambda} \sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} P_k(\chi_{\{r>2^{-k+m}\}}\phi_{\leq k_2+20}(\partial_t + \partial_r)\phi_{k_2}^{(-)\dagger}\phi_{k_3}) \right\|_{L_t^q L_x^p} \\ & \lesssim \sum_{k-10 < k_2 \leq k+m} \sum_{k_3=k_2+O(1)} 2^{(\frac{1}{q}+\frac{2}{p}-1)k} 2^{\lambda k} \sum_{\ell \in \mathbb{Z}} \left\| \chi_{\{r \sim 2^\ell\}} r^{+\lambda}(\partial_t + \partial_r)\phi_{k_2}^{(-)} \right\|_{L_t^q L_x^p} \left\| \chi_{\{r>2^{-k+m}\}}\phi_{k_3} \right\|_{L_t^\infty L_x^\infty} \\ & \lesssim 2^{-\frac{1}{2}m} \sum_{k-10 < k_2 \leq k+m} 2^{(\frac{1}{q}+\frac{2}{p}-\frac{1}{2}+\lambda)(k-k_2)} \|\phi_{k_2}\|_{S_{k_2}}^2 \end{aligned}$$

and in view of the admissible ranges for (q, p) and λ , upon square-summing over $k \in \mathbb{Z}$, we may bound the whole expression by $2^{-\delta m} m^{\frac{1}{2}} \|\phi\|_{S[I]}^2$ for some small $\delta > 0$, which is of the desired form. We proceed analogously with the Y_k norm. When $(\partial_t + \partial_r)$ hits the other inputs, the Z_k^+ norm can be bounded similarly. This finishes the proof of (6.11) and hence the proof of (6.10) for the term (6.6).

In order to show (6.10) for the term (6.7) we can proceed analogously only that we have to introduce an additional splitting

$$\begin{aligned} \sum_{k_2 > k+m} \sum_{k_3=k_2+O(1)} P_k(\phi_{\leq k_2+20}\phi_{k_2}^\dagger\phi_{k_3}) &= \sum_{k_2 > k+m} \sum_{k_3=k_2+O(1)} P_k(\chi_{\{r \leq 2^{-k}\}}\phi_{\leq k_2+20}\phi_{k_2}^\dagger\phi_{k_3}) \\ &+ \sum_{k_2 > k+m} \sum_{k_3=k_2+O(1)} P_k(\chi_{\{r > 2^{-k}\}}\phi_{\leq k_2+20}\phi_{k_2}^\dagger\phi_{k_3}). \end{aligned}$$

Suppose here that the input $\phi_{k_2}^\dagger$ is hit by a derivative, then we achieve an exponential gain in $-k_3$ and thus smallness, either by using Bernstein to place ϕ_{k_3} into $L_t^\infty L_x^2$ or by placing ϕ_{k_3} into $L_t^\infty L_x^\infty$ and using Strauss' improved Sobolev embedding (6.12).

Finally, the proof of (6.10) for the terms (6.8) and (6.9) works similarly to the above estimates and is left to the reader. This finishes the first part of the proof of Proposition 6.3, namely establishing the bound (6.2).

We now turn to the proof of the estimate (6.3). This step is slightly compounded by the fact that the renormalized variable $U_{\leq k-10}\bar{\phi}_k$ is only approximately frequency localized to frequency 2^k up to exponentially decaying tails. We will first show via a divisibility argument that we can partition the time interval $I = \cup_j I_j$ into $N(K, E)$ consecutive intervals $I_j = [t_j, t_{j+1}]$ satisfying

$$(6.13) \quad \left(\sum_{k \in \mathbb{Z}} \left(\sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k-k_0|} \|P_{k_0}(U_{\leq k-10}\bar{\phi}_k)\|_{S_{k_0}[I_j]} \right)^2 \right)^{\frac{1}{2}} \lesssim C(E)$$

for any $0 \leq \alpha < 1$. Then we will infer the desired bound

$$\left(\sum_{k \in \mathbb{Z}} \|\bar{\phi}_k\|_{S_k[I_j]}^2 \right)^{\frac{1}{2}} \lesssim C(E)$$

by decomposing the product $\bar{\phi}_k = U_{\leq k-10}^\dagger (U_{\leq k-10}\bar{\phi}_k)$ into a Littlewood-Paley trichotomy and invoking the bounds (6.13) as well as by introducing a modified version of the gauge transformations $U_{\leq h}$.

We begin with the proof of (6.13). Using the energy estimate (3.1) we find that

$$(6.14) \quad \left(\sum_{k \in \mathbb{Z}} \left(\sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k-k_0|} \|P_{k_0}(U_{\leq k-10}\bar{\phi}_k)\|_{S_{k_0}[I_j]} \right)^2 \right)^{\frac{1}{2}} \lesssim \left(\sum_{k \in \mathbb{Z}} \left(\sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k-k_0|} \|\nabla_{t,x} P_{k_0}(U_{\leq k-10}\bar{\phi}_k)(t_j)\|_{L_x^2} \right)^2 \right)^{\frac{1}{2}} \\ + \left(\sum_{k \in \mathbb{Z}} \left(\sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k-k_0|} \|P_{k_0} \square (U_{\leq k-10}\bar{\phi}_k)\|_{L_t^1 L_x^2[I_j]} \right)^2 \right)^{\frac{1}{2}}.$$

For the first term on the right-hand side of (6.14), we now show in detail that it is bounded by $C(E)$ in the case when the derivative $\nabla_{t,x}$ falls on $\bar{\phi}_k$. The other case when it hits $U_{\leq k-10}$ can be treated similarly. Using Bernstein's inequality and the sharp localization of $\bar{\phi}_k$ to frequency 2^k , we obtain that

$$(6.15) \quad \sum_{k \in \mathbb{Z}} \left(\sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k-k_0|} \|P_{k_0}(U_{\leq k-10}\nabla_{t,x}\bar{\phi}_k)\|_{L_x^2} \right)^2 \lesssim \sum_{k \in \mathbb{Z}} \left(\sum_{k_0 \leq k-10} 2^{\alpha(k-k_0)} \|P_{k_0}((P_{k+O(1)}U_{\leq k-10})\nabla_{t,x}\bar{\phi}_k)\|_{L_x^2} \right)^2 \\ + \sum_{k \in \mathbb{Z}} \|P_{k+O(1)}(U_{\leq k-10}\nabla_{t,x}\bar{\phi}_k)\|_{L_x^2}^2 \\ + \sum_{k \in \mathbb{Z}} \left(\sum_{k_0 > k+10} 2^{\alpha(k_0-k)} \|P_{k_0}((P_{k_0+O(1)}U_{\leq k-10})\nabla_{t,x}\bar{\phi}_k)\|_{L_x^2} \right)^2 \\ \lesssim \sum_{k \in \mathbb{Z}} \left(\sum_{k_0 \leq k-10} 2^{\alpha(k-k_0)} 2^{k_0-2k} \|\nabla_x^2 U_{\leq k-10}\|_{L_x^2} \|\nabla_{t,x}\bar{\phi}_k\|_{L_x^2} \right)^2 \\ + \sum_{k \in \mathbb{Z}} \|\nabla_{t,x}\bar{\phi}_k\|_{L_x^2}^2 \\ + \sum_{k \in \mathbb{Z}} \left(\sum_{k_0 > k+10} 2^{\alpha(k_0-k)} 2^{k-2k_0} \|\nabla_x^2 U_{\leq k-10}\|_{L_x^2} \|\nabla_{t,x}\bar{\phi}_k\|_{L_x^2} \right)^2.$$

Then the bounds on the gauge transformation from Lemma 5.3, the kinetic energy bounds on $\bar{\phi}_k$ from Lemma 5.2 and the fact that $\alpha < 1$ yield that the right-hand side of the previous line is bounded by

$$C(E) \left(\sum_{k_0 \leq k-10} 2^{(\alpha-1)(k-k_0)} + \sum_{k_0 > k+10} 2^{(\alpha-2)(k_0-k)} \right) \sum_{k \in \mathbb{Z}} \|\nabla_{t,x}\bar{\phi}_k\|_{L_x^2}^2 \lesssim C(E).$$

Next we prove by divisibility that on suitable time intervals I_j the second term on the right-hand side of (6.14)

$$(6.16) \quad \left(\sum_{k \in \mathbb{Z}} \left(\sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k-k_0|} \|P_{k_0} \square(U_{\leq k-10} \bar{\phi}_k)\|_{L_t^1 L_x^2[I_j]} \right)^2 \right)^{\frac{1}{2}}$$

is also bounded by $C(E)$. Here we first recall from (5.31) that the wave equation for the renormalized variable $U_{\leq k-10} \bar{\phi}_k$ is given by

$$(6.17) \quad \begin{aligned} \square(U_{\leq k-10} \bar{\phi}_k) &= U_{\leq k-10} (\square \bar{\phi}_k - 2A_{\alpha; \leq k-10} \partial^\alpha \phi_k) \\ &\quad + 2(\partial_\alpha U_{\leq k-10} + U_{\leq k-10} A_{\alpha; \leq k-10}) \partial^\alpha \phi_k \\ &\quad + 2\partial_\alpha U_{\leq k-10} \partial^\alpha (\bar{\phi}_k - \phi_k) \\ &\quad + (\square U_{\leq k-10}) \bar{\phi}_k \end{aligned}$$

and that in Proposition 5.1 we had carefully uncovered the schematic form of each term in the nonlinearity on the right-hand side of (6.17). In view of these identities we observe that we can treat the exponential tails in (6.16) by playing out Bernstein's inequality and the bounds on the gauge transformations from Lemma 5.3, analogously to how we proceeded in the estimate (6.15). It therefore suffices in the following to prove by divisibility that we can partition the time interval $I = \cup_j I_j$ into $N(K, E)$ consecutive intervals I_j satisfying

$$(6.18) \quad \sum_{k \in \mathbb{Z}} \|\square(U_{\leq k-10} \bar{\phi}_k)\|_{L_t^1 L_x^2[I_j]}^2 \lesssim C(E).$$

We now prove this bound separately for each type of term appearing on the right-hand side of the wave equation (6.17) for the renormalized variable $U_{\leq k-10} \bar{\phi}_k$. We start with the term $(\square \bar{\phi}_k - 2A_{\alpha; \leq k-10} \partial^\alpha \phi_k)$ and recall its basic decomposition (5.32). First though, we have to deal with the two extra terms (6.4) and (6.5) that appear due to our modified definition of $\bar{\phi}$ in this proof. For the term (6.4) we use Bernstein's inequality to bound

$$\begin{aligned} &\sum_{k \in \mathbb{Z}} \left\| \sum_{k-10 < k_2 \leq k+m} \sum_{k_3 = k_2 + O(1)} P_k(\chi_{\{r \leq 2^{-k+m}\}} \phi_{\leq k_2+20} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_{k_3}) \right\|_{L_t^1 L_x^2}^2 \\ &\lesssim \sum_{k \in \mathbb{Z}} \left(\sum_{k-10 < k_2 \leq k+m} \sum_{\ell \leq -k+m} 2^{k+\ell} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_{k_2}\|_{L_t^2 L_x^2}^2 \right)^2 \\ &\lesssim \|\phi\|_S \left(\sum_{k \in \mathbb{Z}} \sum_{k-10 < k_2 \leq k+m} \sum_{\ell \leq -k+m} 2^{k+\ell} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_{k_2}\|_{L_t^2 L_x^2}^2 \right). \end{aligned}$$

Then we observe that the last factor in the previous line has the divisibility property and satisfies

$$\sum_{k \in \mathbb{Z}} \sum_{k-10 < k_2 \leq k+m} \sum_{\ell \leq -k+m} 2^{k+\ell} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_{k_2}\|_{L_t^2 L_x^2}^2 \lesssim C(m) \sum_{k \in \mathbb{Z}} \|\phi_k\|_{S_k}^2 \lesssim C(m) \|\phi\|_S^2.$$

It can thus be made sufficiently small when restricted to suitable time intervals noting that at this point the size of the parameter $m \gg 1$ has already been fixed. Similarly, we can gain smallness for the other additional term (6.5) on suitable time intervals. We continue with the first term on the right-hand side of (5.32). Here

we have to distinguish two cases

$$\begin{aligned} & \sum_{k \in \mathbb{Z}} \left\| \sum_{k_2 \leq k-10} \sum_{k_3=k+O(1)} P_k L(\phi_{k_2-10 < \cdot \leq k-10}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \right\|_{L_t^1 L_x^2}^2 \\ & \lesssim \sum_{k \in \mathbb{Z}} \left\| \sum_{k_2 \leq k-10} \sum_{k_3=k+O(1)} P_k \left(\chi_{\{r \leq 2^{-k_2}\}} L(\phi_{k_2-10 < \cdot \leq k-10}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \right) \right\|_{L_t^1 L_x^2}^2 \\ & \quad + \sum_{k \in \mathbb{Z}} \left\| \sum_{k_2 \leq k-10} \sum_{k_3=k+O(1)} P_k \left(\chi_{\{r > 2^{-k_2}\}} L(\phi_{k_2-10 < \cdot \leq k-10}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \right) \right\|_{L_t^1 L_x^2}^2. \end{aligned}$$

In the first case we bound

$$\begin{aligned} & \sum_{k \in \mathbb{Z}} \left\| \sum_{k_2 \leq k-10} \sum_{k_3=k+O(1)} P_k \left(\chi_{\{r \leq 2^{-k_2}\}} L(\phi_{k_2-10 < \cdot \leq k-10}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \right) \right\|_{L_t^1 L_x^2}^2 \\ & \lesssim \sum_{k \in \mathbb{Z}} \left\| \sum_{k_2 \leq k-10} \sum_{k_1 \geq k_2-10} \sum_{\ell \leq -k_1} \chi_{\{r \sim 2^\ell\}} r^{+\frac{1}{2}} |\phi_{k_1}| \|\nabla_{t,x} \phi_{k_2}\| \right\|_{L_t^2 L_x^\infty}^2 \sup_{\ell \in \mathbb{Z}} \left\| \chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_k \right\|_{L_t^2 L_x^2}^2 \\ & \lesssim \left\| \sum_{k_2 \in \mathbb{Z}} \sum_{k_1 \geq k_2-10} \sum_{\ell \leq -k_1} \chi_{\{r \sim 2^\ell\}} r^{+\frac{1}{2}} |\phi_{k_1}| \|\nabla_{t,x} \phi_{k_2}\| \right\|_{L_t^2 L_x^\infty}^2 \|\phi\|_S^2 \end{aligned}$$

and note that the first factor on the right-hand side of the previous line has the divisibility property

$$\left\| \sum_{k_2 \in \mathbb{Z}} \sum_{k_1 \geq k_2-10} \sum_{\ell \leq -k_1} \chi_{\{r \sim 2^\ell\}} r^{+\frac{1}{2}} |\phi_{k_1}| \|\nabla_{t,x} \phi_{k_2}\| \right\|_{L_t^2 L_x^\infty} \lesssim \sum_{k_2 \in \mathbb{Z}} \sum_{k_1 \geq k_2-10} 2^{-\frac{1}{2}k_1} \|\phi_{k_1}\|_{L_t^4 L_x^\infty} \|\nabla_{t,x} \phi_{k_2}\|_{L_t^4 L_x^\infty} \lesssim \|\phi\|_S^2$$

and thus yields smallness on suitable time intervals. In the second case we use the trilinear estimate (4.4) to bound

$$\begin{aligned} & \sum_{k \in \mathbb{Z}} \left\| \sum_{k_2 \leq k-10} \sum_{k_3=k+O(1)} P_k \left(\chi_{\{r > 2^{-k_2}\}} L(\phi_{k_2-10 < \cdot \leq k-10}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \right) \right\|_{L_t^1 L_x^2}^2 \\ & \lesssim \sum_{k \in \mathbb{Z}} \left(\sum_{k_2 \leq k-10} \sum_{k_1 \geq k_2-10} 2^{-\frac{1}{4}(k_1-k_2)} 2^{\frac{1}{4}k_1} \left\| \chi_{\{r > 2^{-k_1}\}} r^{-\frac{1}{4}} \phi_{k_1} \right\|_{L_t^2 L_x^\infty} \|\phi_{k_2}\|_{S_{k_2}} \|\phi_{k_3}\|_{S_{k_3}} \right)^2 \\ & \lesssim \left(\sum_{k_1 \in \mathbb{Z}} 2^{\frac{1}{2}k_1} \left\| \chi_{\{r > 2^{-k_1}\}} r^{-\frac{1}{4}} \phi_{k_1} \right\|_{L_t^2 L_x^\infty}^2 \right) \|\phi\|_S^4 \end{aligned}$$

and then the first factor on the right-hand side has the divisibility property to gain the desired smallness. Next we turn to the second term on the right-hand side of (5.32) and seek to achieve smallness for the expression

$$\sum_{k \in \mathbb{Z}} \left\| \sum_{k_2 > k-10} \sum_{k_1=k_2+O(1)} \sum_{k_3 \leq k_2+O(1)} P_k L(\phi_{k_1}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \right\|_{L_t^1 L_x^2}^2.$$

To this end we note that by Bernstein's inequality and the null form estimate (4.1) it holds that

$$\begin{aligned} & \left\| \sum_{k_2 > k-10} \sum_{k_1=k_2+O(1)} \sum_{k_3 \leq k_2+O(1)} P_k L(\phi_{k_1}, \partial_\alpha \phi_{k_2}, \partial^\alpha \phi_{k_3}) \right\|_{L_t^1 L_x^2}^2 \\ & \lesssim \sum_{k_2 > k-10} \sum_{k_1=k_2+O(1)} \sum_{k_3 \leq k_2+O(1)} 2^{\frac{1}{5}k_1} \|\phi_{k_1}\|_{L_t^{\frac{5}{2}} L_x^\infty} \|\partial_\alpha \phi_{k_2} \partial^\alpha \phi_{k_3}\|_{L_t^{\frac{5}{3}} L_x^{\frac{5}{3}}} \\ & \lesssim \sum_{k_2 > k-10} \sum_{k_3 \leq k_2+O(1)} 2^{-\frac{1}{5}(k_2-k)} 2^{-\frac{1}{10}(k_2-k_3)} \|\phi_{k_2}\|_{S_{k_2}}^2 \|\phi_{k_3}\|_{S_{k_3}}. \end{aligned}$$

Hence, we can exploit the exponential gains in the frequency differences to achieve smallness when the inputs have disparate frequency supports and reduce to the case where all inputs are at frequency $2^{k+O(M)}$

where $M \gg 1$ is chosen sufficiently large depending on $\|\phi\|_S$. But then we obtain smallness by divisibility as in the previous step. The third term on the right-hand side of (5.32) can be treated in exactly the same manner and all quintilinear terms in the fourth term $Q_k(\phi)$ on the right-hand side of (5.32) are easier to control using just Strichartz estimates and the null form estimate (4.1). We leave the details to the reader.

We are left to consider all remaining terms on the right-hand side of the wave equation (6.17) for the renormalized variable $U_{\leq k-10}\bar{\phi}_k$. But in view of the schematic identities (5.33)–(5.35) from Proposition 5.1, these are either of the same type as the first term on the right-hand side of (5.32), which we have already dealt with above, or they are quintilinear and therefore easier to control.

Thus, we can partition the time interval $I = \cup_j I_j$ into $N(K, E)$ consecutive intervals I_j satisfying

$$(6.19) \quad \left(\sum_{k \in \mathbb{Z}} \left(\sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k-k_0|} \|P_{k_0}(U_{\leq k-10}\bar{\phi}_k)\|_{S_{k_0}[I_j]} \right)^2 \right)^{\frac{1}{2}} \lesssim C(E)$$

for any $0 \leq \alpha < 1$. It now remains to transfer these bounds to $\bar{\phi}$, i.e. to show that on each such interval I_j we have that

$$\left(\sum_{k \in \mathbb{Z}} \|\bar{\phi}_k\|_{S_k[I_j]}^2 \right)^{\frac{1}{2}} \lesssim C(E).$$

To this end we use the exact orthogonality of the gauge transformations to write

$$\bar{\phi}_k = U_{\leq k-10}^\dagger U_{\leq k-10} \bar{\phi}_k$$

and then decompose into a Littlewood-Paley trichotomy in view of the localization of $\bar{\phi}_k$ to frequency 2^k ,

$$(6.20) \quad \begin{aligned} \bar{\phi}_k &= P_{\leq k-C}(U_{\leq k-10}^\dagger) P_{k+O(1)}(U_{\leq k-10} \bar{\phi}_k) \\ &\quad + P_{k+O(1)}(U_{\leq k-10}^\dagger) P_{\leq k-C}(U_{\leq k-10} \bar{\phi}_k) \\ &\quad + \sum_{k_1 > k+O(1)} \sum_{k_2 = k_1+O(1)} P_k(P_{k_1}(U_{\leq k-10}^\dagger) P_{k_2}(U_{\leq k-10} \bar{\phi}_k)) \\ &\equiv (LH)_k + (HL)_k + (HH)_k. \end{aligned}$$

We now estimate (6.20) separately for each part of the S_k norm. Here we observe that for those parts of the S_k norm that do not involve a derivative the desired bound follows immediately from (6.19) and the boundedness of $P_{k_1} U_{\leq k_2}^\dagger$ for arbitrary $k_1, k_2 \in \mathbb{Z}$ thanks to the exact orthogonality of $U_{\leq k_2}^\dagger$. Similarly, for all other parts of the S_k norm when the derivative falls on the $U_{\leq k-10} \bar{\phi}_k$ term, the desired bound follows immediately.

Let us therefore begin with the treatment of the Strichartz component of the S_k norm when the derivative $\nabla_{t,x}$ falls onto the term $U_{\leq k-10}^\dagger$. Noting that the gauge transformation bounds from Lemma 5.3 also hold for the transpose $U_{\leq k-10}^\dagger$, we find for the $(LH)_k$ part that

$$\begin{aligned} &2^{(\frac{1}{q} + \frac{2}{p} - 1)k} \sum_{k_1 \leq k-C} \|P_{k_1} \nabla_{t,x} U_{\leq k-10}^\dagger\|_{L_t^\infty L_x^\infty[I_j]} \|P_{k+O(1)}(U_{\leq k-10} \bar{\phi}_k)\|_{L_t^q L_x^p[I_j]} \\ &\lesssim \sum_{k_1 \leq k+O(1)} 2^{-k} 2^{\frac{1}{2}k_1} \|P_{k_1} \nabla_{t,x} U_{\leq k-10}^\dagger\|_{L_t^\infty L_x^4[I_j]} \|P_k(U_{\leq k-10} \bar{\phi}_k)\|_{S_k[I_j]} \\ &\lesssim \sum_{k_1 \leq k+O(1)} 2^{-k} 2^{\frac{1}{2}k_1} 2^{\frac{1}{2}k} C(E) \|P_k(U_{\leq k-10} \bar{\phi}_k)\|_{S_k[I_j]} \\ &\lesssim C(E) \|P_k(U_{\leq k-10} \bar{\phi}_k)\|_{S_k[I_j]}. \end{aligned}$$

Square-summing over $k \in \mathbb{Z}$ and invoking the previously established bound (6.19), we obtain the desired estimate. We proceed analogously with the Strichartz component for the $(HL)_k$ and $(HH)_k$ parts. The weighted $L_t^2 L_x^\infty$ component and the local energy decay component of the S_k norm can also be treated similarly so that

we now turn to the more delicate Z_k^+ norm, the Z_k^- norm being handled in exactly the same manner. We consider in detail the high-high part $(HH)_k$ for the Z_k^+ norm, the $(LH)_k$ and $(HL)_k$ parts being similar.

Here it turns out that we cannot avoid to pick up factors of $\|\phi\|_{S[I_j]}$ which would destroy our final goal to obtain a bound just in terms of the energy E on the interval I_j . For this reason we have to slightly modify the definition of the gauge transformations to introduce another source of smallness to compensate factors of $\|\phi\|_{S[I_j]}$. At this point we recall that the non-perturbative term (5.2),

$$(6.21) \quad 2 \sum_{k_2 \leq k-10} \phi_{\leq k_2-10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_k,$$

in the decomposition (5.2)–(5.7) of the wave maps nonlinearity at frequency 2^k necessitated the introduction of the gauge transformations $U_{\leq h}$. However, letting $m \gg 1$ be a sufficiently large integer depending only on the size of the $S[I]$ norm of ϕ , we can in fact still “peel off” the good term

$$2 \sum_{k-m < k_2 \leq k-10} \chi_{\{r \leq 2^{-k+m}\}} \phi_{\leq k_2-10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_k$$

from (6.21). The latter term can be easily seen to have the divisibility property and can therefore be dealt with accordingly in the previous parts of this proof of Proposition 6.3. Hence, we only have to use a gauge transformation that suitably renormalizes the expression

$$2 \sum_{k_2 \leq k-m} \phi_{\leq k_2-10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_k + 2 \sum_{k-m < k_2 \leq k-10} \chi_{\{r > 2^{-k+m}\}} \phi_{\leq k_2-10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_k.$$

Correspondingly, we modify the definition of the matrix B_h in the construction of the gauge transformations $U_{\leq h}$ in (5.27) to

$$(6.22) \quad \begin{aligned} B_h := & \frac{d}{dh} \sum_{\ell \in \mathbb{Z}} \eta(h+10-m-\ell) (\phi_{\leq \ell-10} \phi_\ell^\dagger - \phi_\ell \phi_{\leq \ell-10}^\dagger) \\ & + \frac{d}{dh} \sum_{\ell \in \mathbb{Z}} (\eta(h-\ell) - \eta(h+10-m-\ell)) \chi_{\{r > 2^{-h-10+m}\}} (\phi_{\leq \ell-10} \phi_\ell^\dagger - \phi_\ell \phi_{\leq \ell-10}^\dagger). \end{aligned}$$

Then we may conclude as in the proof of Proposition 5.1 that $(\partial_t + \partial_r)U_{\leq k-10}^\dagger$ is of the schematic form

$$\begin{aligned} (\partial_t + \partial_r)U_{\leq k-10}^\dagger = & \sum_{k_3 \leq k-m} U_{\leq k_3-1} (\partial_t + \partial_r) L(\phi_{\leq k_3-10}, \phi_{k_3}) \\ & + \sum_{k-m < k_3 \leq k-10} U_{\leq k_3-1} (\partial_t + \partial_r) \chi_{\{r > 2^{-k_3+m}\}} L(\phi_{\leq k_3-10}, \phi_{k_3}). \end{aligned}$$

We now estimate in detail the Z_k^+ norm of the $(HH)_k$ part of (6.20) when $(\partial_t + \partial_r)$ falls onto ϕ_{k_3} , the other case being similar. Thus, we end up having to estimate the Z_k^+ norm of the following schematic expression

$$(6.23) \quad \begin{aligned} & \sum_{k_1 > k+O(1)} \sum_{k_2=k_1+O(1)} \sum_{k_3 \leq k-m} P_k \left(P_{k_1} \left(\sum_{k_3 \leq k-m} U_{\leq k_3-1} (\partial_t + \partial_r) \phi_{k_3} \right) P_{k_2} (U_{\leq k-10} \bar{\phi}_k) \right) \\ & + \sum_{k_1 > k+O(1)} \sum_{k_2=k_1+O(1)} \sum_{k-m < k_3 \leq k-10} P_k \left(P_{k_1} \left(U_{\leq k_3-1} \chi_{\{r > 2^{-k_3+m}\}} (\partial_t + \partial_r) \phi_{k_3} \right) P_{k_2} (U_{\leq k-10} \bar{\phi}_k) \right) \end{aligned}$$

and we begin with the first summand in (6.23). We distinguish the cases $r \leq 2^{-k}$ and $r > 2^{-k}$. When $r \leq 2^{-k}$ we can place the whole output into the $(-)$ component of the Z_k^+ norm using just Bernstein and Strichartz

estimates. Indeed, for the weighted $L_t^q L_x^p$ norm of the $(-)$ component we have

$$\begin{aligned}
& 2^{(\frac{1}{q} + \frac{2}{p} - 1)k} 2^{\lambda k} \sum_{\ell \leq -k} \sum_{k_1 > k + O(1)} \sum_{k_2 = k_1 + O(1)} \sum_{k_3 \leq k - m} \left\| \chi_{\{r \sim 2^\ell\}} r^\lambda P_k \left(P_{k_1} \left(\sum_{k_3 \leq k - m} U_{\leq k_3 - 1} (\partial_t + \partial_r) \phi_{k_3} \right) P_{k_2} (U_{\leq k - 10} \bar{\phi}_k) \right) \right\|_{L_t^q L_x^p [I_j]} \\
& \lesssim 2^{(\frac{1}{q} + \frac{2}{p} - 1)k} \sum_{k_2 > k + O(1)} \sum_{k_3 \leq k - m} \|(\partial_t + \partial_r) \phi_{k_3}\|_{L_t^\infty L_x^p [I_j]} \|P_{k_2} (U_{\leq k - 10} \bar{\phi}_k)\|_{L_t^q L_x^\infty [I_j]} \\
& \lesssim \left(\sum_{k_3 \leq k - m} 2^{(1 - \frac{2}{p})(k_3 - k)} \|\phi_{k_3}\|_{S_{k_3} [I_j]} \right) \sum_{k_2 > k + O(1)} 2^{\frac{1}{q}(k - k_2)} \|P_{k_2} (U_{\leq k - 10} \bar{\phi}_k)\|_{S_{k_2} [I_j]} \\
& \lesssim 2^{-(1 - \frac{2}{p})m} \|\phi\|_{S [I]} \sum_{k_2 > k + O(1)} 2^{\frac{1}{q}(k - k_2)} \|P_{k_2} (U_{\leq k - 10} \bar{\phi}_k)\|_{S_{k_2} [I_j]}.
\end{aligned}$$

Choosing $m \gg 1$ sufficiently large depending on the size of the $S [I]$ norm of ϕ to compensate the factor of $\|\phi\|_{S [I]}$ on the right-hand side of the last line, we may then easily square-sum over $k \in \mathbb{Z}$ and invoke the estimate (6.19) to obtain a final bound just in terms of the energy E . Similarly we bound the Y_k norm of the $(-)$ component.

Let us now consider the case when $r > 2^{-k}$. Here we distinguish whether $(\partial_t + \partial_r) \phi_{k_3}$ is of $(+)$ type or of $(-)$ type. If it is of $(+)$ type, we place the whole expression into the $(+)$ component of the Z_k^+ norm. To this end we consider the $L_{t+r}^2 L_{t-r}^\infty$ part of the $(+)$ component. For $0 < \lambda < \frac{1}{2}$ we have

$$\begin{aligned}
& 2^{-(\frac{1}{2} - \lambda)k} \left\| r^\lambda \chi_{\{r > 2^{-k}\}} \sum_{k_1 > k + O(1)} \sum_{k_2 = k_1 + O(1)} \sum_{k_3 \leq k - m} P_k \left(P_{k_1} \left(\sum_{k_3 \leq k - m} U_{\leq k_3 - 1} (\partial_t + \partial_r) \phi_{k_3}^{(+)} \right) P_{k_2} (U_{\leq k - 10} \bar{\phi}_k) \right) \right\|_{L_{t+r}^2 L_{t-r}^\infty [I_j]} \\
& \lesssim 2^{-(\frac{1}{2} - \lambda)k} \sum_{k_2 > k + O(1)} \sum_{k_3 \leq k - m} \|r^\lambda (\partial_t + \partial_r) \phi_{k_3}^{(+)}\|_{L_{t+r}^2 L_{t-r}^\infty [I_j]} \|P_{k_2} (U_{\leq k - 10} \bar{\phi}_k)\|_{L_t^\infty L_x^\infty} \\
& \lesssim \left(\sum_{k_3 \leq k - m} 2^{-(\frac{1}{2} - \lambda)(k - k_3)} \|\phi_{k_3}\|_{S_{k_3} [I]} \right) \sum_{k_2 \in \mathbb{Z}} \|P_{k_2} (U_{\leq k - 10} \bar{\phi}_k)\|_{S_{k_2} [I_j]} \\
& \lesssim 2^{-(\frac{1}{2} - \lambda)m} \|\phi\|_{S [I]} \sum_{k_2 \in \mathbb{Z}} \|P_{k_2} (U_{\leq k - 10} \bar{\phi}_k)\|_{S_{k_2} [I_j]}
\end{aligned}$$

and upon choosing $m \gg 1$ sufficiently large, we may easily square-sum in $k \in \mathbb{Z}$ and obtain a final bound just in terms of the energy E thanks to (6.19). For $\lambda = \frac{1}{2}$ we unfortunately do not straightaway have summability in $k_3 \leq k - m$. In this case we split the weight $r^{\frac{1}{2}} = r^{\frac{1}{2} - \delta} r^\delta$ for some small $\delta > 0$ and absorb r^δ into the high-frequency factor $P_{k_2} (U_{\leq k - 10} \bar{\phi}_k)$ via Strauss' improved Sobolev embedding in the radial case

$$\left\| \chi_{\{r > 2^{-k}\}} r^\delta P_{k_2} (U_{\leq k - 10} \bar{\phi}_k) \right\|_{L_t^\infty L_x^\infty [I_j]} \lesssim 2^{(\frac{1}{2} - \delta)k} 2^{-\frac{1}{2}k_2} \|P_{k_2} (U_{\leq k - 10} \bar{\phi}_k)\|_{L_t^\infty L_x^2 [I_j]}.$$

Then we can again sum over $k_3 \leq k - m$ and gain a smallness factor $2^{-\delta m}$. The Y_k norm of the $(-)$ component of the Z_k^+ norm is easier to bound. It therefore remains to consider the case when $(\partial_t + \partial_r) \phi_{k_3}$ is of $(-)$ type. Here we try to place the whole expression into the $(-)$ component of the Z_k^+ norm. We start with the weighted $L_t^q L_x^p$ part

$$2^{(\frac{1}{q} + \frac{2}{p} - 1)k} 2^{\lambda k} \sum_{\ell > -k} \left\| \chi_{\{r \sim 2^\ell\}} r^\lambda \sum_{k_1 > k + O(1)} \sum_{k_2 = k + O(1)} \sum_{k_3 \leq k - m} P_k \left(P_{k_1} \left(\sum_{k_3 \leq k - m} U_{\leq k_3 - 1} (\partial_t + \partial_r) \phi_{k_3}^{(-)} \right) P_{k_2} (U_{\leq k - 10} \bar{\phi}_k) \right) \right\|_{L_t^q L_x^p [I_j]},$$

where we aim to estimate $(\partial_t + \partial_r) \phi_{k_3}^{(-)}$ in $L_t^q L_x^\infty$, while placing the high-frequency factor $P_{k_2} (U_{\leq k - 10} \bar{\phi}_k)$ into $L_t^\infty L_x^p$. In order to ensure summability over the low frequencies $k_3 \leq k - m$, we exploit that we can absorb a weight $r^{\frac{1}{2} - \frac{1}{p}}$ into $P_{k_2} (U_{\leq k - 10} \bar{\phi}_k)$ via the estimate

$$\left\| r^{\frac{1}{2} - \frac{1}{p}} P_{k_2} f \right\|_{L_t^\infty L_x^p} \lesssim 2^{-(\frac{1}{2} + \frac{1}{p})k_2} \|\nabla_x P_{k_2} f\|_{L_x^2},$$

which follows from interpolating Strauss' improved Sobolev embedding with the trivial energy estimate. This yields the following bound on the weighted $L_t^q L_x^p$ part of the $(-)$ component of the Z_k^+ norm

$$\begin{aligned} & \sum_{k_3 \leq k-10} 2^{(\frac{1}{q}-1+\lambda)k} 2^{(1-\lambda+\frac{1}{2}-\frac{1}{p}-\frac{1}{q})k_3} \|\phi_{k_3}\|_{S_{k_3}[I_j]} \sum_{k_2 > k+O(1)} 2^{\frac{2}{p}k} 2^{-(\frac{1}{2}+\frac{1}{p})k_2} \|P_{k_2}(U_{\leq k-10}\bar{\phi}_k)\|_{S_{k_2}[I_j]} \\ & \lesssim 2^{-(1-\lambda+\frac{1}{2}-\frac{1}{p}-\frac{1}{q})m} \sum_{k_2 > k+O(1)} 2^{-(\frac{1}{2}+\frac{1}{p})(k_2-k)} \|P_{k_2}(U_{\leq k-10}\bar{\phi}_k)\|_{S_{k_2}[I_j]}, \end{aligned}$$

which yields a smallness gain $2^{-\delta m}$ on account of the fact that $0 < \lambda < 1$ and $\frac{1}{q} + \frac{1}{p} < \frac{1}{2}$. Then we can easily square-sum this bound over $k \in \mathbb{Z}$ and obtain a final estimate just in terms of E by invoking (6.19). The Y_k norm of the $(-)$ component of the Z_k^+ norm can again be treated easily.

Finally, we have to estimate the Z_k^+ norm of the second summand in (6.23). Here we proceed similarly to the above treatment of the first summand only that we gain smallness using Strauss' improved Sobolev embedding and the restriction $r > 2^{-k+m}$. This finishes the proof of Proposition 6.3. \square

Next, we give the proof of Proposition 6.2 using the weak divisibility of the S norm.

Proof of Proposition 6.2. We may assume that the interval I is of the form $I = [0, T]$ for some $0 < T \leq \infty$. Then we first use Proposition 6.3 to partition $I = \cup_{j=1}^N I_j$ into $N \equiv N(\|\phi\|_{S[I]}, E)$ consecutive intervals $I_j = [t_{j-1}, t_j]$ with $t_0 = 0$ and $t_N = T$ such that $\|\phi\|_{S[I_j]} \leq C(E)$ for $j = 1, \dots, N$, where $C(E) > 0$ is a constant that depends only on the energy E of the wave map ϕ . For each interval $I_j = [t_{j-1}, t_j]$ we introduce a frequency envelope

$$c_k^{(j)} := \sum_{\ell \in \mathbb{Z}} 2^{-\sigma|k-\ell|} \|P_\ell \phi[t_{j-1}]\|_{\dot{H}_x^1 \times L_x^2}, \quad k \in \mathbb{Z},$$

for some small constant $\sigma > 0$, and show via an iterative bootstrap argument that

$$\|P_k \phi\|_{S_k[I_j]} \lesssim c_k^{(j)},$$

which then implies the assertion of the proposition. To this end we further partition each interval $I_j = \cup_i I_{ji}$ into finitely many consecutive intervals I_{ji} which have suitable divisibility properties. On every interval I_{ji} we now run a bootstrap argument. Starting with I_{j1} we make the bootstrap assumption

$$\|P_k \phi\|_{S_k[I_{j1}]} \leq C c_k^{(j)}$$

for some sufficiently large, absolute constant $C > 0$ and now show that this implies the improved bound

$$\|P_k \phi\|_{S_k[I_{j1}]} \leq \frac{C}{2} c_k^{(j)}.$$

Then we continue analogously on all remaining intervals I_{ji} . Here we again have to use the modified definition (6.6) of $\bar{\phi}_k$ and the modified definition (6.22) of the gauge transformations $U_{\leq h}$ to have an additional source of smallness at our disposal as in the preceding proof of Proposition 6.3. Then it is easy to show that for any $0 \leq \alpha < 1$,

$$\sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k-k_0|} \|P_{k_0} \nabla_{t,x}(U_{\leq k-10}\bar{\phi}_k(t_j))\|_{L_x^2} \ll C c_k^{(j)}.$$

Moreover, picking the intervals I_{ji} suitably, we may infer by divisibility arguments that for any $0 \leq \alpha < 1$,

$$\sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k-k_0|} \|P_{k_0} \square(U_{\leq k-10}\bar{\phi}_k)\|_{L_t^1 L_x^2[I_{j1}]} \ll C c_k^{(j)}$$

and then the energy estimate (3.1) implies that

$$\sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k-k_0|} \|P_{k_0}(U_{\leq k-10}\bar{\phi}_k)\|_{S_{k_0}[I_{j1}]} \ll C c_k^{(j)}.$$

Finally, we argue similarly as in the preceding proof of Proposition 6.3 to pass from the last bound back to ϕ_k and recover the improved bound

$$\|P_k \phi\|_{S_k[I_{j1}]} \leq \frac{C}{2} c_k^{(j)}.$$

This finishes the proof of Proposition 6.2. \square

We are now in a position to provide the proof of the main result of this section.

Proof of Proposition 6.1. Suppose that $I \neq \mathbb{R}$. Since we have $\|\phi\|_{S[I]} < \infty$ by assumption, Proposition 6.2 yields frequency envelope bounds for the evolution of ϕ on its maximal time interval of existence I ,

$$\|P_k \phi\|_{S_k[I]} \leq C c_k,$$

where $\{c_k\}_{k \in \mathbb{Z}}$ is a frequency envelope covering the initial data $\phi[0]$. Then it is a standard argument to infer that a subcritical norm of $\phi[t]$ must stay finite on I , i.e.

$$\sup_{t \in I} \|\phi[t]\|_{H_x^s \times H_x^{s-1}} < \infty$$

for some $s > 1$. But then the local well-posedness theory [8–10, 12, 13] implies that the evolution of ϕ extends smoothly beyond the time interval I , contradicting the maximality of I . Thus, we must have $I = \mathbb{R}$ and it remains to prove the scattering assertion. To this end we first note that in the gauged wave equation

$$\square(U_{\leq k-10} \bar{\phi}_k) \equiv F_k, \quad k \in \mathbb{Z},$$

we have that $\|F_k\|_{L_t^1 L_x^2[\mathbb{R}]} \lesssim c_k$. Now for $k \ll -1$ or $k \gg 1$ we already know from the a priori bounds that $\|\nabla_{t,x} \phi_k\|_{L_t^\infty L_x^2} \lesssim c_k$ has very small norm, so it suffices to consider $k = O(1)$. Picking a sufficiently large time $T > 0$ such that

$$\|F_k\|_{L_t^1 L_x^2([T, \infty))} \ll 1,$$

we then obtain that for $k = O(1)$,

$$U_{\leq k-10} \bar{\phi}_k(t) = S(t-T)(U_{\leq k-10} \bar{\phi}_k[T]) + o_{L_t^\infty \dot{H}_x^1([T, \infty))}(1) \quad \text{for } t > T.$$

In order to infer scattering, it then suffices to show that

$$\lim_{t \rightarrow \infty} \|\nabla_{t,x}(U_{\leq k-10} \bar{\phi}_k(t) - \phi_k(t))\|_{L_x^2} = 0,$$

for which in turn it suffices to prove that $\lim_{t \rightarrow \infty} \|U_{\leq k-10}(t) - Id\|_{L_x^\infty} = 0$ as well as $\lim_{t \rightarrow \infty} \|\phi_k(t)\|_{L_x^\infty} = 0$. On the one hand, it is easy to see that

$$\lim_{t \rightarrow \infty} \|S(t-T)(U_{\leq k-10} \bar{\phi}_k[T])\|_{L_x^\infty} = 0,$$

and the localization to frequency $k = O(1)$ also implies that the error $o_{L_t^\infty \dot{H}_x^1([T, \infty))}(1)$ goes to zero in L_x^∞ . This then implies that $\bar{\phi}_k(t)$ converges to zero in L_x^∞ . To pass from here to ϕ_k again requires modifying the definition of $\bar{\phi}_k$ as in the previous proof of Proposition 6.3 so that the difference $\bar{\phi}_k - \phi_k$ becomes arbitrarily small. Having uniform smallness for ϕ_k in L_x^∞ , we then obtain that

$$\lim_{t \rightarrow \infty} \|U_{\leq k-10}(t) - Id\|_{L_x^\infty} = 0,$$

which implies

$$\bar{\phi}_k(t) = S(t-T)(U_{\leq k-10} \bar{\phi}_k[T]) + o_{L_t^\infty \dot{H}_x^1([T, \infty))}(1) \quad \text{for } t > T.$$

Then the modified definition of $\bar{\phi}_k$ yields that

$$\phi_k(t) = S(t-T)(U_{\leq k-10} \bar{\phi}_k[T]) + o_{L_t^\infty \dot{H}_x^1([T, \infty))}(1) \quad \text{for } t > T,$$

which gives scattering. \square

7. CONCENTRATION COMPACTNESS STEP

In this section we begin with the actual proof of Theorem 1.1. We recall that our goal is to show that there exists a non-decreasing function $K: [0, \infty) \rightarrow [0, \infty)$ with the following property: Let $(\phi_0, \phi_1): \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ be radially symmetric, classical initial data of energy E . Then there exists a unique, classical, global wave map $\phi: \mathbb{R}^{1+2} \rightarrow \mathbb{S}^m$ with initial data $\phi[0] = (\phi_0, \phi_1)$ satisfying the a priori bound

$$\|\phi\|_S \leq K(E).$$

Once we have established this a priori bound, the scattering assertion of Theorem 1.1 is an immediate consequence of Proposition 6.1.

We argue by contradiction and assume that Theorem 1.1 fails. Then the existence of the function $K(\cdot)$ yielding the a priori bounds must fail at some finite energy level. Correspondingly, the following set of energies must be non-empty

$$\mathcal{E} := \left\{ E : \sup_{\{\phi : E[\phi] \leq E\}} \|\phi\|_{S[I]} = +\infty \right\},$$

where the supremum is taken over all radial, classical wave maps $\phi: I \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$ defined on some time interval I and with energy $E[\phi] \leq E$. By the small energy global regularity result from Theorem 5.4, the infimum of the set \mathcal{E} has to be strictly positive

$$E_{crit} := \inf \mathcal{E} > 0.$$

Thus, we may pick a sequence of radially symmetric, classical wave maps $\phi^n: I^n \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$, $n \geq 1$, with maximal intervals of existence I^n such that

$$\lim_{n \rightarrow \infty} E[\phi^n] = E_{crit}, \quad \lim_{n \rightarrow \infty} \|\phi^n\|_{S[I^n]} = +\infty.$$

In the following we call such a sequence of wave maps *essentially singular*. Moreover, we denote the associated essentially singular sequence of initial data $\{\phi^n[0]\}_{n \geq 1}$ by $\{(\phi_0^n, \phi_1^n)\}_{n \geq 1}$. Here, the subscripts in (ϕ_0^n, ϕ_1^n) should not to be confused with frequency localizations, but this will always be clear from the context.

Our goal is now to rule out the existence of such an essentially singular sequence of wave maps $\{\phi^n\}_{n \geq 1}$, hence proving Theorem 1.1. To this end we follow the general philosophy of the concentration compactness/rigidity method introduced by Kenig-Merle [6, 7], but more precisely we shall follow the implementation of this strategy for energy critical wave maps into the hyperbolic plane as in [17] as well as for the energy critical Maxwell-Klein-Gordon equation as in [16]. In this section we carry out a “twisted” Bahouri-Gérard type profile decomposition that takes into account the strong low-high interactions in the wave maps nonlinearity. It enables us to extract from the essentially singular sequence a non-trivial minimal blowup solution to (WM) whose orbit satisfies a strong compactness property. Then we exclude the existence of such a minimal blowup solution in the rigidity argument of the next, and final, section of this paper.

The first step consists in decomposing the essentially singular sequence of data $\{\phi^n[0]\}_{n \geq 1}$ into frequency atoms using the Métivier-Schochet procedure [21] as in Bahouri-Gérard [1]. Roughly speaking, the basic idea then goes as follows. Ultimately, we would like to conclude that upon passing to a subsequence, if necessary, the essentially singular sequence of data $\{\phi^n[0]\}_{n \geq 1}$ consists of exactly one frequency atom which in turn consists of exactly one concentration profile (to be defined precisely in Subsection 7.4) of asymptotic energy E_{crit} . In this scenario, the sequence $\{\phi^n[0]\}_{n \geq 1}$ has sufficient compactness properties that allow us to pass to a certain limit whose wave maps evolution will be the desired minimal blowup solution to (WM) as detailed in Subsection 7.5. In order to rule out all other possible scenarios, we seek to prove uniform in n , finite, global S norm bounds on the sequence of wave maps evolutions $\{\phi^n\}_{n \geq 1}$, which would contradict that the sequence is essentially singular. To this end we first achieve control over the wave maps evolutions of certain low frequency truncations of the essentially singular sequence of data $\{\phi^n[0]\}_{n \geq 1}$. Using a finite inductive procedure over the increasing size of the frequency supports of these low frequency truncations, we then conclude uniform in n , finite, global S norm bounds on the actual essentially singular sequence

$\{\phi^n\}_{n \geq 1}$. This inductive procedure over the increasing size of the frequency supports also enables us to disentangle the strong low-high frequency interactions in the wave maps nonlinearity.

7.1. Decomposition into frequency atoms. We now turn to the details of the decomposition of the essentially singular sequence of data $\{\phi^n[0]\}_{n \geq 1}$ into frequency atoms. Here we follow relatively closely Section 9.1 and Section 9.2 in [17] as well as Section 7.2 in [16], which in turn partially mimic Section III.1 in Bahouri-Gérard [1]. First, we need to introduce some terminology from [1].

We call a sequence of positive numbers $\{\lambda_n\}_{n \geq 1}$ a *scale*. Two scales $\{\lambda_n^a\}_{n \geq 1}$ and $\{\lambda_n^b\}_{n \geq 1}$ are *orthogonal* if

$$\lim_{n \rightarrow \infty} \frac{\lambda_n^a}{\lambda_n^b} + \frac{\lambda_n^b}{\lambda_n^a} = +\infty.$$

Let $\{(f^n, g^n)\}_{n \geq 1}$ be a bounded sequence of functions in $\dot{H}_x^1(\mathbb{R}^2) \times L_x^2(\mathbb{R}^2)$ and let $\{\lambda_n\}_{n \geq 1}$ be a scale. Then the sequence $\{(f^n, g^n)\}_{n \geq 1}$ is called λ_n -*oscillatory* if

$$\lim_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \left(\int_{\{\lambda_n |\xi| \leq \frac{1}{R}\}} |\widehat{\nabla_x f^n}(\xi)|^2 + |\widehat{g^n}(\xi)|^2 d\xi + \int_{\{\lambda_n |\xi| \geq R\}} |\widehat{\nabla_x f^n}(\xi)|^2 + |\widehat{g^n}(\xi)|^2 d\xi \right) = 0$$

and we say that the sequence $\{(f^n, g^n)\}_{n \geq 1}$ is λ_n -*singular* if for all $0 < a < b$,

$$\lim_{n \rightarrow \infty} \int_{\{a \leq \lambda_n |\xi| \leq b\}} |\widehat{\nabla_x f^n}(\xi)|^2 + |\widehat{g^n}(\xi)|^2 d\xi = 0.$$

Using the Métivier-Schochet procedure [21] as implemented in Bahouri-Gérard [1], we start off with a decomposition

$$\phi^n[0] = \sum_{a=1}^{\Lambda} \tilde{\phi}^{na}[0] + \tilde{\phi}^{n\Lambda}[0]$$

of the essentially singular sequence of data $\{\phi^n[0]\}_{n \geq 1}$ into λ_n^a -oscillatory *frequency atoms* $\tilde{\phi}^{na}[0]$, $1 \leq a \leq \Lambda$, for pairwise orthogonal frequency scales $\{\lambda_n^a\}_{n \geq 1}$ and into a remainder term $\tilde{\phi}^{n\Lambda}[0]$, which is λ_n^a -singular for $1 \leq a \leq \Lambda$ and has the smallness property

$$\lim_{\Lambda \rightarrow \infty} \limsup_{n \rightarrow \infty} \|\tilde{\phi}^{n\Lambda}[0]\|_{\dot{B}_{2,\infty}^1 \times \dot{B}_{2,\infty}^0} = 0.$$

Next we achieve a sharp frequency localization of the atoms by picking a sequence $R_n \rightarrow \infty$ growing sufficiently slowly and by then setting

$$\phi^{na}[0] := P_{[\mu_n^a - \log R_n, \mu_n^a + \log R_n]} \tilde{\phi}^{na}[0] \quad \text{for } a = 1, \dots, \Lambda$$

and

$$\phi^{n\Lambda}[0] := P_{\cap_{a=1}^{\Lambda} [\mu_n^a - \log R_n, \mu_n^a + \log R_n]^c} \tilde{\phi}^{n\Lambda}[0],$$

where we use the notation $\mu_n^a = -\log(\lambda_n^a)$. Then we have the new frequency atom decomposition

$$(7.1) \quad \phi^n[0] = \sum_{a=1}^{\Lambda} \phi^{na}[0] + \phi^{n\Lambda}[0]$$

with the same properties as above, but now with the additional sharp frequency localization property of the atoms. Moreover, we may assume that the atoms are ordered in terms of the increasing size of their frequency support scales $(\lambda_n^a)^{-1}$.

As described above, ultimately we would like to conclude that there is exactly one atom in the decomposition (7.1), i.e. $\Lambda = 1$, which is of asymptotic energy E_{crit} . If this is the case, we proceed directly as in Subsection 7.4. Otherwise, we now conclude via a finite inductive procedure that the sequence $\{\phi^n\}_{n \geq 1}$

cannot be essentially singular. To this end we fix an integer Λ_0 sufficiently large such that upon passing to a subsequence, if necessary,

$$\sum_{a > \Lambda_0} \limsup_{n \rightarrow \infty} \|\phi^{na}[0]\|_{H_x^1 \times L_x^2}^2 \leq \varepsilon_0,$$

where $\varepsilon_0 > 0$ is a sufficiently small constant that plays the role of a perturbative threshold in the key bootstrap argument in Proposition 7.9 in Subsection 7.3. In particular, ε_0 will be chosen sufficiently small depending only on the size of E_{crit} and it will be chosen to be less than the small energy global regularity threshold from Theorem 5.4. Then we observe that due to the sharp frequency localizations of the atoms $\phi^{na}[0]$, $1 \leq a \leq \Lambda_0$, the remainder term $\phi^{n\Lambda_0}[0]$ gets split into $\Lambda_0 + 1$ “frequency shells”

$$(7.2) \quad \phi^{n\Lambda_0}[0] = \phi^{n\Lambda_0^{(0)}}[0] + \phi^{n\Lambda_0^{(1)}}[0] + \dots + \phi^{n\Lambda_0^{(\Lambda_0)}}[0],$$

where $\phi^{n\Lambda_0^{(0)}}[0]$ shall denote the lowest frequency component.

Our first step now consists in showing that the lowest frequency “non-atomic” component $\phi^{n\Lambda_0^{(0)}}[0]$ can be globally evolved and satisfies finite S norm bounds just in terms of E_{crit} uniformly for all sufficiently large n . Since the component $\phi^{n\Lambda_0^{(0)}}[0]$ may still have large energy, in order to be able to infer these S norm bounds by bootstrap, we split $\phi^{n\Lambda_0^{(0)}}[0]$ into finitely many pieces

$$\phi^{n\Lambda_0^{(0)}}[0] = \sum_{\ell=1}^{C_1} P_{J_\ell} \phi^{n\Lambda_0^{(0)}}[0]$$

by means of frequency localization to consecutive frequency intervals $J_\ell = [a_\ell, b_\ell]$ such that

$$(-\infty, \mu_n^1 - \log R_n] = \cup_{\ell=1}^{C_1} J_\ell$$

and such that for $\ell = 1, \dots, C_1$,

$$\|P_{J_\ell} \phi^{n\Lambda_0^{(0)}}[0]\|_{H_x^1 \times L_x^2}^2 \lesssim \varepsilon_0.$$

Here we recall that $(-\infty, \mu_n^1 - \log R_n]$ is the (dyadic) frequency support of $\phi^{n\Lambda_0^{(0)}}[0]$. The number C_1 of such frequency intervals is of the order $O(\frac{E_{crit}}{\varepsilon_0})$ since $\|\phi^{n\Lambda_0^{(0)}}[0]\|_{H_x^1 \times L_x^2}^2 \lesssim E_{crit}$.

We now intend to inductively show that if

$$\sum_{\ell \leq L} P_{J_\ell} \phi^{n\Lambda_0^{(0)}}[0] = P_{\leq b_L} \phi^{n\Lambda_0^{(0)}}[0]$$

can be globally evolved with uniform finite S norm bounds in terms of E_{crit} for some $L \geq 1$ (for all sufficiently large n), then we may also conclude this for

$$\sum_{\ell \leq L+1} P_{J_\ell} \phi^{n\Lambda_0^{(0)}}[0] = P_{\leq b_{L+1}} \phi^{n\Lambda_0^{(0)}}[0],$$

and the induction start for the first piece $P_{J_1} \phi^{n\Lambda_0^{(0)}}[0]$ is provided by the small energy global regularity result from Theorem 5.4.

However, up to this point we have totally ignored that the frequency truncations $P_{\leq b_L} \phi^{n\Lambda_0^{(0)}}[0] \equiv P_{\leq b_L} \phi^n[0]$ are not “geometric” in the sense that they are not actual maps $\mathbb{R}^2 \rightarrow T\mathbb{S}^m$ into the unit sphere and therefore do not constitute suitable initial data for the wave maps equation (WM). To overcome this issue we just project the frequency truncations back to the sphere, using the normal projection operator. As we shall see next, this operation is well-defined and the frequency localization properties are approximately preserved up to exponential tails if around the frequency cut-offs a certain Besov norm smallness condition is satisfied.

We begin with a general lemma that shows that the frequency truncations $P_{\leq \alpha_n} \phi_0^n$ of maps $\phi_0^n: \mathbb{R}^2 \rightarrow \mathbb{S}^m$, $n \geq 1$, remain close to the sphere if around the frequency cut-offs $\{\alpha_n\}_{n \geq 1}$ a certain Besov norm smallness holds.

Lemma 7.1. *Let $\{\alpha_n\}_{n \geq 1}$ be a sequence of real numbers. Given $\delta_1 > 0$ one can choose $\delta_0 \ll \delta_1$ sufficiently small (depending on E_{crit}) so that if we have*

$$(7.3) \quad \limsup_{n \rightarrow \infty} \|P_{[\alpha_n, \alpha_n + M]} \phi_0^n\|_{\dot{B}_{2, \infty}^1} \leq \delta_0$$

for $M \sim \log((1 + E_{crit})\delta_1^{-1})$, then it holds that

$$\limsup_{n \rightarrow \infty} \|1 - (P_{\leq \alpha_n} \phi_0^n)^\dagger (P_{\leq \alpha_n} \phi_0^n)\|_{L_x^\infty} \lesssim \delta_1.$$

Proof. By definition we have

$$(P_{\leq \alpha_n} \phi_0^n + P_{> \alpha_n} \phi_0^n)^\dagger (P_{\leq \alpha_n} \phi_0^n + P_{> \alpha_n} \phi_0^n) = 1$$

and therefore

$$1 - (P_{\leq \alpha_n} \phi_0^n)^\dagger (P_{\leq \alpha_n} \phi_0^n) = 2P_{\leq \alpha_n + 10}((P_{\leq \alpha_n} \phi_0^n)^\dagger (P_{> \alpha_n} \phi_0^n)) + P_{\leq \alpha_n + 10}((P_{> \alpha_n} \phi_0^n)^\dagger (P_{> \alpha_n} \phi_0^n)).$$

Then we use Bernstein's inequality and (7.3) to bound

$$\begin{aligned} \limsup_{n \rightarrow \infty} \|P_{\leq \alpha_n + 10}((P_{\leq \alpha_n} \phi_0^n)^\dagger (P_{> \alpha_n} \phi_0^n))\|_{L_x^\infty} &\lesssim \limsup_{n \rightarrow \infty} \|P_{\leq \alpha_n} \phi_0^n\|_{L_x^\infty} \sum_{k=\alpha_n}^{\alpha_n + 15} 2^k \|P_k \phi_0^n\|_{L_x^2} \\ &\lesssim \limsup_{n \rightarrow \infty} \sum_{k=\alpha_n}^{\alpha_n + 15} \|\nabla_x P_k \phi_0^n\|_{L_x^2} \\ &\lesssim \delta_0 \lesssim \delta_1. \end{aligned}$$

Similarly, we obtain

$$\begin{aligned} \limsup_{n \rightarrow \infty} \|P_{\leq \alpha_n + 10}((P_{> \alpha_n} \phi_0^n)^\dagger (P_{> \alpha_n} \phi_0^n))\|_{L_x^\infty} &\lesssim \limsup_{n \rightarrow \infty} \sum_{k \geq \alpha_n} 2^{2\alpha_n} \|P_k \phi_0^n\|_{L_x^2}^2 \\ &\lesssim \limsup_{n \rightarrow \infty} \sum_{k=\alpha_n}^{\alpha_n + M} 2^{2\alpha_n - 2k} \|\nabla_x P_k \phi_0^n\|_{L_x^2}^2 + \sum_{k > \alpha_n + M} 2^{-2M} \|\nabla_x P_k \phi_0^n\|_{L_x^2}^2 \\ &\lesssim M\delta_0 + 2^{-2M} E_{crit} \\ &\lesssim \delta_1, \end{aligned}$$

where in the last step we use that $M \sim \log((1 + E_{crit})\delta_1^{-1})$ and choose $\delta_0 \lesssim \delta_1 \log^{-1}((1 + E_{crit})\delta_1^{-1})$. \square

For suitably chosen frequency cut-offs α_n , the frequency truncations $P_{\leq \alpha_n} \phi_0^n$ therefore stay close to the sphere. In order to recover exact maps into the sphere, we then simply project the frequency truncations $P_{\leq \alpha_n} \phi_0^n$ back to the sphere, using the normal projection operator Π . Hence, we set

$$(7.4) \quad \Pi_{\leq \alpha_n} \phi_0^n := \Pi(P_{\leq \alpha_n} \phi_0^n) := \frac{P_{\leq \alpha_n} \phi_0^n}{|P_{\leq \alpha_n} \phi_0^n|}.$$

For the associated frequency truncations $P_{\leq \alpha_n} \phi_1^n$ of the time derivatives ϕ_1^n in our essentially singular sequence of data $\phi^n[0] = (\phi_0^n, \phi_1^n)$, we just use linear orthogonal projection onto the respective fiber of the tangent space and map $P_{\leq \alpha_n} \phi_1^n$ to

$$(7.5) \quad \Pi_{\leq \alpha_n} \phi_1^n := P_{\leq \alpha_n} \phi_1^n - \langle \Pi_{\leq \alpha_n} \phi_0^n, P_{\leq \alpha_n} \phi_1^n \rangle \Pi_{\leq \alpha_n} \phi_0^n \in T_{\Pi_{\leq \alpha_n} \phi_0^n} \mathbb{S}^m.$$

Thus, after frequency truncation of the data $\phi^n[0] = (\phi_0^n, \phi_1^n): \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ to dyadic frequencies less than α_n , we may recover the data pair

$$\Pi_{\leq \alpha_n} \phi^n[0] := (\Pi_{\leq \alpha_n} \phi_0^n, \Pi_{\leq \alpha_n} \phi_1^n): \mathbb{R}^2 \rightarrow T\mathbb{S}^m.$$

In the next proposition we prove that this operation approximately preserves the frequency localization if the data satisfy a Besov norm smallness condition around the frequency cut-offs.

Proposition 7.2. Let $\{\alpha_n\}_{n \geq 1}$ be a sequence of real numbers. Given $\delta_1 > 0$ one can choose $\delta_0 \ll \delta_1$ sufficiently small depending on the size of E_{crit} so that if we have

$$(7.6) \quad \limsup_{n \rightarrow \infty} \|P_{[\alpha_n - M, \alpha_n + M]} \phi^n[0]\|_{\dot{B}_{2,\infty}^1 \times \dot{B}_{2,\infty}^0} \leq \delta_0$$

for $M \sim \log((1 + E_{crit})\delta_1^{-1})$, then for all sufficiently large n it holds that

$$(7.7) \quad \|P_k(P_{\leq \alpha_n} \phi^n[0] - \Pi_{\leq \alpha_n} \phi^n[0])\|_{\dot{H}_x^1 \times L_x^2} \lesssim \delta_1 2^{-|k - \alpha_n|} \quad \text{for } k \in \mathbb{Z}.$$

Proof. We begin with the proof of (7.7) for the component $\Pi_{\leq \alpha_n} \phi_0^n$ for the low frequencies $k \leq \alpha_n$. Using that $\phi_0^n = \Pi(\phi_0^n)$, we may write

$$\phi_0^n - \Pi(P_{\leq \alpha_n} \phi_0^n) = \int_{\alpha_n}^{\infty} \frac{d}{dh} (\Pi(P_{\leq h} \phi_0^n)) dh = \int_{\alpha_n}^{\infty} (P_h \phi_0^n) \Pi'(P_{\leq h} \phi_0^n) dh.$$

Thus, we obtain from Bernstein's inequality that

$$\begin{aligned} & \|P_k \nabla_x (P_{\leq \alpha_n} \phi_0^n - \Pi_{\leq \alpha_n} \phi_0^n)\|_{L_x^2} \\ & \lesssim \left\| P_k \int_{\alpha_n}^{\infty} (\nabla_x P_h \phi_0^n) \Pi'(P_{\leq h} \phi_0^n) dh \right\|_{L_x^2} + \left\| P_k \int_{\alpha_n}^{\infty} (P_h \phi_0^n) (\nabla_x P_{\leq h} \phi_0^n) \Pi''(P_{\leq h} \phi_0^n) dh \right\|_{L_x^2} \\ & \lesssim 2^k \int_{\alpha_n}^{\infty} \|\nabla_x P_h \phi_0^n\|_{L_x^2} 2^{-h} \|\nabla_x P_{\leq h-10} \Pi'(P_{\leq h} \phi_0^n)\|_{L_x^2} dh \\ & \quad + 2^k \int_{\alpha_n}^{\infty} 2^{-h} \|\nabla_x P_h \phi_0^n\|_{L_x^2} \|\nabla_x P_{\leq h} \phi_0^n\|_{L_x^2} \|\Pi''(P_{\leq h} \phi_0^n)\|_{L_x^\infty} dh \\ & \lesssim 2^k \int_{\alpha_n}^{\infty} 2^{-h} \|\nabla_x P_h \phi_0^n\|_{L_x^2} \|\nabla_x P_{\leq h} \phi_0^n\|_{L_x^2} \|\Pi''(P_{\leq h} \phi_0^n)\|_{L_x^\infty} dh. \end{aligned}$$

Using (7.6), we conclude that for all sufficiently large n , the previous line is bounded by

$$\begin{aligned} & 2^k E_{crit}^{\frac{1}{2}} \int_{\alpha_n}^{\alpha_n + M} 2^{-h} \|\nabla_x P_h \phi_0^n\|_{L_x^2} dh + 2^k E_{crit}^{\frac{1}{2}} \int_{\alpha_n + M}^{\infty} 2^{-h} \|\nabla_x P_h \phi_0^n\|_{L_x^2} dh \\ & \lesssim 2^{k - \alpha_n} E_{crit}^{\frac{1}{2}} M \delta_0 + 2^{k - \alpha_n} E_{crit} 2^{-M}. \end{aligned}$$

For the high frequencies $k > \alpha_n$, we use that

$$\nabla_x P_k \Pi(P_{\leq \alpha_n} \phi_0^n) = \nabla_x P_k \int_{-\infty}^{\alpha_n} (P_h \phi_0^n) \Pi'(P_{\leq h} \phi_0^n) dh.$$

Then we have

$$\begin{aligned} (7.8) \quad & \|\nabla_x P_k \Pi(P_{\leq \alpha_n} \phi_0^n)\|_{L_x^2} \lesssim 2^{-k} \|\nabla_x^2 P_k \Pi(P_{\leq \alpha_n} \phi_0^n)\|_{L_x^2} \\ & \lesssim 2^{-k} \left\| P_k \int_{-\infty}^{\alpha_n} (\nabla_x^2 P_h \phi_0^n) \Pi'(P_{\leq h} \phi_0^n) dh \right\|_{L_x^2} \\ & \quad + 2^{-k} \left\| P_k \int_{-\infty}^{\alpha_n} (\nabla_x P_h \phi_0^n) (\nabla_x P_{\leq h} \phi_0^n) \Pi''(P_{\leq h} \phi_0^n) dh \right\|_{L_x^2}. \end{aligned}$$

We bound the first term on the right-hand side of (7.8) by

$$\begin{aligned} & 2^{-k} \int_{-\infty}^{\alpha_n} 2^h \|\nabla_x P_h \phi_0^n\|_{L_x^2} dh \lesssim 2^{-k} \int_{\alpha_n - M}^{\alpha_n} 2^h \|\nabla_x P_h \phi_0^n\|_{L_x^2} dh + 2^{-k} \int_{-\infty}^{\alpha_n - M} 2^h \|\nabla_x P_h \phi_0^n\|_{L_x^2} dh \\ & \lesssim 2^{-(k - \alpha_n)} (M \delta_0 + E_{crit}^{\frac{1}{2}} 2^{-M}), \end{aligned}$$

while the second term on the right-hand side of (7.8) can be estimated by

$$\begin{aligned}
2^{-k} \int_{-\infty}^{\alpha_n} \|\nabla_x P_h \phi_0^n\|_{L_x^2} \|\nabla_x P_{\leq h} \phi_0^n\|_{L_x^\infty} dh &\lesssim 2^{-k} \int_{\alpha_n-M}^{\alpha_n} \|\nabla_x P_h \phi_0^n\|_{L_x^2} 2^h \|\nabla_x P_{\leq h} \phi_0^n\|_{L_x^2} dh \\
&\quad + 2^{-k} \int_{-\infty}^{\alpha_n-M} \|\nabla_x P_h \phi_0^n\|_{L_x^2} 2^h \|\nabla_x P_{\leq h} \phi_0^n\|_{L_x^2} dh \\
&\lesssim 2^{-(k-\alpha_n)} (1 + E_{crit}) (M\delta_0 + 2^{-M}).
\end{aligned}$$

Thus, we have inferred that

$$\left\| \nabla_x P_k (P_{\leq \alpha_n} \phi_0^n - \Pi_{\leq \alpha_n} \phi_0^n) \right\|_{L_x^2} \lesssim (1 + E_{crit}) (M\delta_0 + 2^{-M}) 2^{-|k-\alpha_n|}.$$

Since $M \sim \log((1 + E_{crit})\delta_1^{-1})$, we obtain (7.7) for the component $\Pi_{\leq \alpha_n} \phi_0^n$ upon choosing $\delta_0 \ll \delta_1$ sufficiently small depending on the size of E_{crit} .

Finally, we turn to the proof of (7.7) for the time derivative component $\Pi_{\leq \alpha_n} \phi_1^n$. For the low frequencies $k \leq \alpha_n$ we exploit the fact that $(\phi_0^n)^\dagger \phi_1^n = 0$ to write

$$\begin{aligned}
P_{\leq \alpha_n} \phi_1^n - \Pi_{\leq \alpha_n} \phi_1^n &= -\langle \Pi_{\leq \alpha_n} \phi_0^n, P_{\leq \alpha_n} \phi_1^n \rangle \Pi_{\leq \alpha_n} \phi_0^n \\
&= \int_{\alpha_n}^{\infty} \frac{d}{dh} \langle \Pi_{\leq h} \phi_0^n, P_{\leq h} \phi_1^n \rangle \Pi_{\leq h} \phi_0^n dh \\
&= \int_{\alpha_n}^{\infty} \langle \Pi_{\leq h} \phi_0^n, P_{\leq h} \phi_1^n \rangle (P_h \phi_0^n) \Pi'(P_{\leq h} \phi_0^n) dh \\
&\quad + \int_{\alpha_n}^{\infty} \langle (P_h \phi_0^n) \Pi'(P_{\leq h} \phi_0^n), P_{\leq h} \phi_1^n \rangle \Pi(P_{\leq h} \phi_0^n) dh \\
&\quad + \int_{\alpha_n}^{\infty} \langle \Pi_{\leq h} \phi_0^n, P_h \phi_1^n \rangle \Pi_{\leq h} \phi_0^n dh.
\end{aligned} \tag{7.9}$$

Then we may bound the first term on the right-hand side of (7.9) by

$$\begin{aligned}
&\left\| P_k \int_{\alpha_n}^{\infty} \langle \Pi_{\leq h} \phi_0^n, P_{\leq h} \phi_1^n \rangle (P_h \phi_0^n) \Pi'(P_{\leq h} \phi_0^n) dh \right\|_{L_x^2} \\
&\lesssim 2^k \int_{\alpha_n}^{\infty} \|\Pi_{\leq h} \phi_0^n\|_{L_x^\infty} \|P_{\leq h} \phi_1^n\|_{L_x^2} \|P_h \phi_0^n\|_{L_x^2} \|\Pi'(P_{\leq h} \phi_0^n)\|_{L_x^\infty} dh \\
&\lesssim 2^k E_{crit}^{\frac{1}{2}} \int_{\alpha_n}^{\infty} 2^{-h} \|\nabla_x P_h \phi_0^n\|_{L_x^2} dh \\
&\lesssim 2^{k-\alpha_n} (1 + E_{crit}) (M\delta_0 + 2^{-M}),
\end{aligned}$$

which is of the desired form, while the other two terms on the right-hand side of (7.9) can be dealt with similarly. Proving (7.7) for $\Pi_{\leq \alpha_n} \phi_1^n$ for the high frequencies $k > \alpha_n$ is a variant of the previous estimates. \square

Remark 7.3. We note that Proposition 7.2 and its proof are reminiscent of Proposition 11.1 in Sterbenz-Tataru [29] where it is shown that for initial data sets with small “energy dispersion” frequency truncation followed by normal projection approximately preserves the frequency localization properties up to exponentially decaying tails.

Thus, for the ensuing induction on frequency process, we shall use $\Pi_{\leq b_L} \phi^{n\Lambda_0^{(0)}}[0]$ as the actual data for the wave maps evolution. But in order for this data $\Pi_{\leq b_L} \phi^{n\Lambda_0^{(0)}}[0]$ to approximately have the same frequency localization properties up to exponentially decaying tails as the frequency truncations $P_{\leq b_L} \phi^{n\Lambda_0^{(0)}}[0]$, Proposition 7.2 requires the Besov norm smallness condition (7.6). This, in particular, forces us to implement the following delicate selection procedure for the endpoints of the intervals $J_\ell = [a_\ell, b_\ell]$.

We first use the Métivier-Schochet procedure [21] to carry out a further refined decomposition of the lowest frequency “non-atomic” component $\phi^{n\Lambda_0^{(0)}}[0]$ into finitely many “smaller” frequency atoms. More precisely, given any $\delta_0 > 0$ we decompose into

$$\phi^{n\Lambda_0^{(0)}}[0] = \sum_{j=1}^{J(\Lambda_0^{(0)})} \phi^{na_j^{(0)}}[0] + \phi^{nJ(\Lambda_0^{(0)})}[0]$$

such that the “small” atoms $\phi^{na_j^{(0)}}[0]$ are frequency localized to

$$|\xi| \in [(\lambda_n^{a_j^{(0)}})^{-1}(R_n^{(0)})^{-1}, (\lambda_n^{a_j^{(0)}})^{-1}R_n^{(0)}]$$

for a sequence $R_n^{(0)} \rightarrow \infty$ growing sufficiently slowly as $n \rightarrow \infty$ and such that

$$(7.10) \quad \limsup_{n \rightarrow \infty} \|\phi^{nJ(\Lambda_0^{(0)})}[0]\|_{\dot{B}_{2,\infty}^1 \times \dot{B}_{2,\infty}^0} \leq \delta_0.$$

By the orthogonality of the frequency scales, we may assume that for sufficiently large n the frequency intervals

$$[(\lambda_n^{a_j^{(0)}})^{-1}(R_n^{(0)})^{-1}, (\lambda_n^{a_j^{(0)}})^{-1}R_n^{(0)}]$$

are disjoint for different $a_j^{(0)}$.

Then we are in the position to describe how the frequency intervals J_ℓ can be chosen inductively. Assume that the intervals

$$J_1 = [a_1, b_1], \dots, J_{\ell-1} = [a_{\ell-1}, b_{\ell-1}]$$

have already been specified. Then we pick the maximal interval $[a_\ell, \tilde{b}_\ell]$ with the property that $a_\ell = b_{\ell-1}$ and such that

$$\|P_{[a_\ell, \tilde{b}_\ell]} \phi^{n\Lambda_0^{(0)}}[0]\|_{\dot{H}_x^1 \times L_x^2}^2 = \varepsilon_0.$$

If it happens that

$$\tilde{b}_\ell \in [-\log(\lambda_n^{a_j^{(0)}}) - 2\log(R_n^{(0)}), -\log(\lambda_n^{a_j^{(0)}}) + 2\log(R_n^{(0)})],$$

then we shift the interval endpoint upwards and take

$$b_\ell = -\log(\lambda_n^{a_j^{(0)}}) + 2\log(R_n^{(0)}),$$

otherwise we set $b_\ell = \tilde{b}_\ell$.

The point of this construction is that for $\ell = 1, \dots, C_1$, the frequency intervals $J_\ell = [a_\ell, b_\ell]$ now satisfy

$$\|P_{J_\ell} \phi^{n\Lambda_0^{(0)}}[0]\|_{\dot{H}_x^1 \times L_x^2}^2 \lesssim \varepsilon_0$$

and have the desired *key property* that for any fixed $M > 0$, it holds that

$$(7.11) \quad \limsup_{n \rightarrow \infty} \|P_{[b_\ell - M, b_\ell + M]} \phi^{n\Lambda_0^{(0)}}[0]\|_{\dot{B}_{2,\infty}^1 \times \dot{B}_{2,\infty}^0} \leq \delta_0$$

for all sufficiently large n .

7.2. Interlude: A concept of energy class radial wave maps. At the end of the entire concentration compactness step in this Section 7 we want to extract a minimal blowup solution to the wave maps equation (WM) that is merely of energy class. However, since the local well-posedness theory [8–10, 12, 13] only pertains to data of regularity $H_x^{1+}(\mathbb{R}^2) \times H_x^{0+}(\mathbb{R}^2)$, we first of all have to introduce a notion of the wave maps evolution of radially symmetric energy class data. We shall achieve this analogously to the procedures in [17] and [16] by regularization and reduction to the small energy case via finite speed of propagation. We begin with the following “high-frequency perturbation” lemma.

Lemma 7.4. Let $\phi : [0, T] \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$ be a wave map which is supported in the frequency interval $(-\infty, K]$ for some $K \in \mathbb{R}$ up to an exponentially decaying tail, more precisely such that

$$\|\phi_k\|_{S_k([0, T])} \leq \delta_1 2^{-\sigma(k-K)}, \quad k \geq K,$$

for some positive constants $0 < \delta_1 \ll 1$ and $\sigma > 0$. Suppose that

$$\|\phi\|_{S([0, T])} \leq C_1$$

for some $C_1 > 0$. Then there exists $\delta \equiv \delta(C_1, \sigma)$ such that if $\epsilon[0] \in H_x^{1+} \times H_x^{0+}$ is a data pair such that $(\phi + \epsilon)[0]$ constitutes admissible initial data mapping into $T\mathbb{S}^m$ and such that

$$\|\epsilon[0]\|_{\dot{H}_x^1 \times L_x^2} \leq \delta \quad \text{with } \delta_1 \ll \delta,$$

and moreover if $\epsilon[0]$ is supported up to exponentially decaying tails at frequencies $[K, \infty)$ in the sense that

$$\|P_k \epsilon[0]\|_{\dot{H}_x^1 \times L_x^2} \leq \delta 2^{-\sigma(K-k)}, \quad k < K,$$

then the data $(\phi + \epsilon)[0]$ can be evolved on $[0, T] \times \mathbb{R}^2$ as a wave map $(\phi + \epsilon)(t, x)$ of class $H_x^{1+} \times H_x^{0+}$ and we have that

$$\|\epsilon\|_{S([0, T])} \leq C_2(C_1, \sigma) \delta.$$

Proof. From the basic local existence theory [8–10, 12, 13] we know that the data $(\phi + \epsilon)[0]$ can be evolved locally in time as a wave map of class $H_x^{1+} \times H_x^{0+}$. It remains to show that this evolution extends all the way up to time $t = T$. This we do by a bootstrap argument which is a simplified version of the one that will be used for controlling the lowest frequency “non-atomic” component in Subsection 7.3. In fact, we make the bootstrap assumption

$$\|\epsilon_k\|_{S_k([0, T])} \leq C(c_k + d_k)$$

for some sufficiently large $C > 0$, where $\{c_k\}_{k \in \mathbb{Z}}$ is such that $c_k = 0$ for $k < K$ and $\sum_{k \in \mathbb{Z}} c_k^2 \leq \delta^2$, while $\{d_k\}_{k \in \mathbb{Z}}$ is such that $d_k = \delta 2^{-\sigma|k-K|}$ for $k < K$ and $d_k = 0$ for $k \geq K$. Then we argue as in the bootstrap argument in the proof of Proposition 7.9 in Subsection 7.3. \square

Now assume that we are given a radially symmetric data pair $\phi[0] : \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ of energy class $\phi[0] \in (\dot{H}_x^1 \cap L_x^\infty) \times L_x^2$. Then we can pass to the frequency truncated data $\Pi_{<K} \phi[0]$ for any $K \in \mathbb{R}$. As each of these data sets are of class $H_x^{1+} \times H_x^{0+}$ we have a canonical evolution for them on some local time intervals, which may however a priori depend on K . But then we observe

Lemma 7.5. Let $\phi[0] : \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ be radially symmetric data of energy class $\phi[0] \in (\dot{H}_x^1 \cap L_x^\infty) \times L_x^2$. Then there exists a time $T_* > 0$, $T_* \equiv T_*(\phi[0])$, and some sufficiently large $K_0 \in \mathbb{R}$ such that for any $K \geq K_0$, the wave maps evolution $\phi^{(K)}(t, x)$ of the data $\Pi_{<K} \phi[0]$ exists on $[0, T_*] \times \mathbb{R}^2$, and moreover, we have a uniform bound

$$\|\phi^{(K)}\|_{S([0, T_*])} \leq C(\phi[0]), \quad K \geq K_0.$$

Proof. The claim will follow from the small energy global regularity result from Theorem 5.4 via Huygen’s principle and a simple partition of unity argument to patch together the global-in-space solution from spatially localized ones. To this end we define the annuli

$$C_{R_1, R_2} := \{x \in \mathbb{R}^2 : R_1 \leq |x| \leq R_2\}, \quad 0 \leq R_1 < R_2 \leq +\infty$$

and denote by $\eta > 0$ the cutoff for the small energy global regularity theory. Given radially symmetric data $\phi[0] : \mathbb{R}^2 \rightarrow T\mathbb{S}^m$, there exists a covering of \mathbb{R}^2 by finitely many annuli C_{R_{j-1}, R_j} , $j = 1, 2, \dots, J$, with $R_0 = 0$, $R_J = +\infty$, $R_j \leq 2R_{j-1}$ for $j \in \{2, \dots, J-1\}$, and such that $\phi[0]|_{C_{R_{j-1}, R_j}}$ coincides with the restriction to C_{R_{j-1}, R_j} of some radially symmetric data $\phi^{(j)}[0] : \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ of energy less than η . To see this, pick finitely many such annuli C_{R_{j-1}, R_j} with the property that

$$\|\phi[0]|_{C_{\frac{1}{2}R_{j-1}, 2R_j}}\|_{\dot{H}_x^1 \times L_x^2} \ll \eta.$$

Then observe that for any $x, y \in C_{\frac{1}{2}R_{j-1}, 2R_j}$, $j \in \{2, \dots, J-1\}$, with $r = |x|$ and $\tilde{r} = |y|$, we have that

$$|\phi(x) - \phi(y)| = \left| \int_{\tilde{r}}^r \nabla_r \phi(r) dr \right| \lesssim \left| \log \left(\frac{r}{\tilde{r}} \right) \right|^{\frac{1}{2}} \|\nabla_x \phi\|_{C_{\frac{1}{2}R_{j-1}, 2R_j}} \ll \eta.$$

Then picking $x_0 \in C_{R_{j-1}, R_j}$ arbitrarily and introducing

$$\phi^{(j)} := \Pi \left(\chi_{C_{\frac{1}{2}R_{j-1}, 2R_j}} \phi + (1 - \chi_{C_{\frac{1}{2}R_{j-1}, 2R_j}}) \phi(x_0) \right), \quad j \in \{2, \dots, J-1\},$$

where $\{\chi_{C_{\frac{1}{2}R_{j-1}, 2R_j}}\}_{j=1, \dots, J}$ is a smooth partition of unity of \mathbb{R}^2 with $\chi_{C_{\frac{1}{2}R_{j-1}, 2R_j}}|_{C_{R_{j-1}, R_j}} \equiv 1$, we obtain

$$\|\nabla_x \phi^{(j)}\|_{L_x^2(\mathbb{R}^2)} \ll \eta$$

as well as

$$\phi^{(j)}|_{C_{R_{j-1}, R_j}} \equiv \phi.$$

Further, we set

$$\partial_t \phi^{(j)} := \chi_{C_{\frac{1}{2}R_{j-1}, 2R_j}} \partial_t \phi - \chi_{C_{\frac{1}{2}R_{j-1}, 2R_j}} ((\partial_t \phi)^\dagger \phi^{(j)}) \phi^{(j)}.$$

Then we have $\partial_t \phi^{(j)}|_{C_{R_{j-1}, R_j}} = \partial_t \phi$ and it is easily seen that

$$\|\partial_t \phi^{(j)}\|_{L_x^2(\mathbb{R}^2)} \ll \eta.$$

Importantly, the constructed data pair $\phi^{(j)}[0] = (\phi^{(j)}, \partial_t \phi^{(j)}): \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ is again radially symmetric. It is also straightforward to modify this construction on $C_{\frac{1}{2}R_{j-1}, 2R_j}$ for $j = 1$ and for $j = M$. Replacing $\phi[0]$ by $\phi^{(K)}[0] = \Pi_{<K} \phi[0]$ results in the data $\phi^{(j,K)}[0]$, which satisfy the same bounds (with a slightly different implied constant) for large enough K . It then follows from the small energy global regularity theory that we can evolve these data $\phi^{(j,K)}[0]$ to global wave maps and a simple application of Huygen's principle implies that the evolutions $\phi^{(K)}$ exist on a joint time interval $[0, T_*]$ for some $0 < T_* \ll 1$ for all large enough $K \geq K_0$. Moreover, we may infer uniform bounds

$$\|\phi^{(K)}\|_{S([0, T_*])} \leq C_*(\phi[0]), \quad K \geq K_0.$$

□

Combining the two preceding lemmas, we can now deduce the desired concept of energy class evolution. Given radially symmetric energy class data $\phi[0]: \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ with $\phi[0] \in (\dot{H}_x^1 \cap L_x^\infty) \times L_x^2$, we see that by Lemma 7.5 there exists $T_*(\phi[0]) > 0$ such that for all sufficiently large $K \geq K_0$ the wave maps evolutions $\phi^{(K)}(t, x)$ exist on a joint time interval $[0, T_*]$ and moreover, using Lemma 7.4, we see that the sequence $\{\phi^{(K)}\}_{K \geq K_0}$ converges in the sense of $\|\cdot\|_{S([0, T_*])}$. It also follows that the limit is canonical, i.e. it does not depend on the precise choice of regularization. Correspondingly, we introduce

Definition 7.6. Let $\phi[0]: \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ be a radially symmetric energy class data pair $\phi[0] \in (\dot{H}_x^1 \cap L_x^\infty) \times L_x^2$ and let $\{\phi^{(K)}[0]\}_K$ be a sequence of frequency truncated data $\phi^{(K)}[0] = \Pi_{<K} \phi[0]$ such that $\phi^{(K)}[0] \rightarrow \phi[0]$ as $K \rightarrow \infty$ in the sense of $\dot{H}_x^1 \times L_x^2$. We denote by $\phi^{(K)}$ the smooth local wave maps evolutions of the data $\phi^{(K)}[0]$ and define $I = (-T_0, T_1) = \cup \tilde{I}$ to be the union of all open time intervals $\tilde{I} \ni 0$ such that

$$\sup_{J \subset \tilde{I}, J \text{ closed}} \liminf_{K \rightarrow \infty} \|\phi^{(K)}\|_{S[J]} < \infty.$$

Then we define the wave maps evolution of $\phi[0]$ on $I \times \mathbb{R}^2$ to be

$$\phi[t] := \lim_{K \rightarrow \infty} \phi^{(K)}[t], \quad t \in I,$$

where the limit is taken in the energy topology. We refer to I as the maximal lifespan of ϕ . For any closed time interval $J \subset I$, we set

$$\|\phi\|_{S[J]} := \lim_{K \rightarrow \infty} \|\phi^{(K)}\|_{S[J]}.$$

Our above considerations also imply the following characterization of the maximal lifespan I of an energy class wave maps evolution.

Lemma 7.7. *Let ϕ , $\phi^{(K)}$ and I be as in the preceding Definition 7.6. Suppose that $I \neq \mathbb{R}$. Then it must hold that*

$$\sup_{J \subset I, J \text{ closed}} \liminf_{K \rightarrow \infty} \|\phi^{(K)}\|_{S[J]} = \infty.$$

7.3. Evolving the lowest frequency non-atomic part. In this subsection we prove that the lowest frequency non-atomic part $\phi^{n\Lambda_0^{(0)}}[0]$, more precisely the associated wave map data $\Pi_{\leq \mu_n^1 - \log(R_n)} \phi^n[0]$, can be globally evolved for all sufficiently large n with uniform in n bounds on the S norms of the global evolutions just in terms of E_{crit} . The end result is summarized in the following proposition.

Proposition 7.8. *Let $\phi^{n\Lambda_0^{(0)}}[0]$ be defined as in (7.2). Then provided $\delta_0 \ll \delta_1 \ll \varepsilon_0$ are chosen sufficiently small depending on the size of E_{crit} , the associated wave map initial data $\Pi_{\leq \mu_n^1 - \log(R_n)} \phi^n[0]$ can be evolved globally in time for all sufficiently large n and their evolutions denoted by $\Pi_{\leq \mu_n^1 - \log(R_n)} \phi^n$ obey*

$$\|\Pi_{\leq \mu_n^1 - \log(R_n)} \phi^n\|_{S[\mathbb{R}]} \leq C(E_{crit})$$

uniformly for all sufficiently large n for some constant $C(E_{crit}) > 0$ that depends only on E_{crit} .

We shall prove Proposition 7.8 via a finite inductive procedure over the increasing size of the frequencies, more precisely we shall inductively conclude that for $L = 1, \dots, C_1$, the data $\Pi_{\leq b_L} \phi^n[0]$ can be globally evolved for all sufficiently large n with uniform in n bounds on the S norms of the evolutions. The induction start is given by the small energy global regularity result from Theorem 5.4 and the induction step is provided by the following key proposition of this subsection.

Proposition 7.9. *Assume that for some $1 \leq L < C_1$, the data $\Pi_{\leq b_L} \phi^{n\Lambda_0^{(0)}}[0]$ can be evolved globally in time for all large n and that their evolutions denoted by $\Pi_{\leq b_L} \phi^{n\Lambda_0^{(0)}}$ satisfy*

$$\|\Pi_{\leq b_L} \phi^{n\Lambda_0^{(0)}}\|_{S[\mathbb{R}]} \leq C_2.$$

Then provided $\delta_0 \ll \delta_1$ and $\delta_1 \equiv \delta_1(C_2, E_{crit})$ are chosen sufficiently small, there exists $C_3 \equiv C_3(C_2) > 0$ such that for all sufficiently large n , $\Pi_{\leq b_{L+1}} \phi^{n\Lambda_0^{(0)}}[0]$ can be globally evolved and their evolutions denoted by $\Pi_{\leq b_{L+1}} \phi^{n\Lambda_0^{(0)}}$ satisfy

$$\|\Pi_{\leq b_{L+1}} \phi^{n\Lambda_0^{(0)}}\|_{S[\mathbb{R}]} \leq C_3.$$

Proof. To simplify the notation in the following proof we shall write

$$\phi \equiv \Pi_{\leq b_L} \phi^{n\Lambda_0^{(0)}}, \quad \epsilon \equiv \Pi_{\leq b_{L+1}} \phi^{n\Lambda_0^{(0)}} - \Pi_{\leq b_L} \phi^{n\Lambda_0^{(0)}},$$

keeping in mind that ϕ and ϵ depend on n . Since by assumption ϕ exists globally in time with finite S norm, we have frequency envelope bounds for ϕ thanks to Proposition 6.2. In particular, due to the Besov smallness (7.11) around the frequency interval endpoint b_L , this implies that for all sufficiently large n ,

$$(7.12) \quad \|P_k \phi\|_{S_k[\mathbb{R}]} \lesssim \delta_1 2^{-\sigma(k-b_L)} \text{ for } k > b_L.$$

While ϕ exists globally in time, ϵ only exists locally in time and for now, any statement we make about ϵ is meant locally in time on some interval I_0 around $t = 0$. In order to prove global existence and finite S norm bounds for ϵ , we use Proposition 6.3 to partition the time axis $\mathbb{R} = \cup_{j=1}^N I_j$ into $N \equiv N(C_2, E_{crit})$ many time intervals I_j with the property that

$$\|\phi\|_{S[I_j]} \lesssim C(E_{crit}) \quad \text{for } j = 1, \dots, N.$$

We tacitly assume that these intervals are intersected with I_0 and that the interval I_1 contains time $t = 0$. Our strategy is to iteratively prove S norm bounds for ϵ by bootstrap on each interval I_j . Here we encounter

the danger that the energy of ϵ could keep growing as we move to later time intervals and could thereby leave the perturbative regime before we would have concluded S norm bounds on ϵ on all time intervals I_1, \dots, I_N . However, we will see that the energy transfer between ϵ and ϕ is controlled by the underlying Besov error δ_0 , which therefore implies approximate energy conservation for ϵ . We now fix the interval I_1 with the understanding that all the arguments in this step can be carried out for the later time intervals I_2, \dots, I_N .

For the ensuing bootstrap argument we introduce another smallness parameter $\delta_2 > 0$ satisfying

$$0 < \delta_0 \ll \delta_1 \ll \delta_2 \ll \varepsilon_0 \ll 1.$$

Here we recall that δ_0 is the underlying Besov error (7.10) of the non-atomic component in the frequency atom decomposition and that δ_1 controls the approximate frequency localization (7.7) of the normal projection operator. Then we make the following bootstrap assumptions.

Bootstrap assumptions: Suppose that

$$(7.13) \quad \|P_k \epsilon\|_{S_k[I_1]} \leq C(c_k + d_k) \text{ for } k \in \mathbb{Z},$$

where $C \equiv C(E_{crit}) \gg 1$ is a sufficiently large constant and where $\{c_k\}_{k \in \mathbb{Z}}$ satisfies

$$(7.14) \quad c_k = 0 \text{ for } k < b_L, \quad c_k = 0 \text{ for } k > b_{L+1}, \quad \sum_{k \in \mathbb{Z}} c_k^2 \lesssim \varepsilon_0$$

as well as

$$(7.15) \quad \sum_{b_L \leq k \leq b_L + M} c_k^2 \leq \delta_2^2 \text{ for some } M \gg 1 \text{ with } 2^{-M} \leq \delta_2^3.$$

Moreover, $\{d_k\}_{k \in \mathbb{Z}}$ satisfies

$$(7.16) \quad d_k = \delta_2 2^{-\sigma(b_L - k)} \text{ for } k < b_L, \quad d_k = 0 \text{ for } b_L \leq k \leq b_{L+1}, \quad d_k = \delta_2 2^{-\sigma(k - b_{L+1})} \text{ for } k > b_{L+1}.$$

The idea here is that $\{d_k\}_{k \in \mathbb{Z}}$ incorporates the frequency leakage to ϵ coming from ϕ in the difference equation for ϵ as well as the approximate frequency localization of the data $\epsilon[0]$ up to exponential tails coming from application of the normal projection operator Π . Upon writing

$$(7.17) \quad \epsilon[0] = P_{[b_L, b_{L+1}]} \phi^{n\Lambda_0^{(0)}}[0] + (\Pi_{\leq b_{L+1}} \phi^{n\Lambda_0^{(0)}} - P_{\leq b_{L+1}} \phi^{n\Lambda_0^{(0)}})[0] - (\Pi_{\leq b_L} \phi^{n\Lambda_0^{(0)}} - P_{\leq b_L} \phi^{n\Lambda_0^{(0)}})[0],$$

it becomes clear that by the choice of the frequency intervals $[b_L, b_{L+1}]$, by the Besov smallness (7.11) around the endpoints of these intervals and by the approximate preservation of the frequency localization by the normal projection operator Π as established in Proposition 7.2, we have for all sufficiently large n that

$$\|P_k \epsilon[0]\|_{\dot{H}_x^1 \times L_x^2} \lesssim c_k + d_k.$$

In a first step we use a direct energy conservation argument to recover the envelope for the low frequencies $k < b_L$ with respect to the kinetic energy norm. Then we will crucially use this bound and the “renormalized difference equation” for ϵ to recover the full bound.

Lemma 7.10. *If ϵ satisfies the assumption (7.13), then for some $1 \ll D_0 \ll C$ the following improved bound for the kinetic energy component for the low frequencies $k < b_L$ holds*

$$(7.18) \quad \|\nabla_{t,x} P_k \epsilon\|_{L_t^\infty L_x^2[I_1]} \leq D_0(c_k + d_k).$$

Proof. We may assume that $k = 0$, so in particular we have $b_L > 0$ in the following. Then we observe that ϵ_0 satisfies the wave equation

$$\square \epsilon_0 = -P_0((\phi + \epsilon)\partial_\alpha(\phi + \epsilon)^\dagger \partial^\alpha(\phi + \epsilon)) + P_0(\phi \partial_\alpha \phi^\dagger \partial^\alpha \phi) \equiv P_0 X(\phi, \epsilon).$$

Passing to the corresponding energy identity, we see that we need to show

$$\left| \int_{I_1} \int_{\mathbb{R}^2} P_0 X(\phi, \epsilon)^\dagger \partial_t \epsilon_0 \, dx \, dt \right| \ll C^2 \delta_2^2 2^{-2\sigma b_L}.$$

We now prove this bound separately for each term in

$$(7.19) \quad -X(\phi, \epsilon) = \epsilon \partial_\alpha \phi^\dagger \partial^\alpha \phi + 2\phi \partial_\alpha \epsilon^\dagger \partial^\alpha \phi + \epsilon \partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon + 2\epsilon \partial_\alpha \epsilon^\dagger \partial^\alpha \phi + \phi \partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon.$$

(i) *Contribution of $P_0(\epsilon \partial_\alpha \phi^\dagger \partial^\alpha \phi)$.* We split into

$$(7.20) \quad P_0(\epsilon \partial_\alpha \phi^\dagger \partial^\alpha \phi) = P_0(P_{>b_L} \epsilon \partial_\alpha \phi^\dagger \partial^\alpha \phi) + P_0(P_{[-10, b_L]} \epsilon \partial_\alpha \phi^\dagger \partial^\alpha \phi) + P_0(P_{\leq -10} \epsilon \partial_\alpha \phi^\dagger \partial^\alpha \phi).$$

Then we have for the first term

$$\left| \int_{I_1} \int_{\mathbb{R}^2} P_0(P_{>b_L} \epsilon \partial_\alpha \phi^\dagger \partial^\alpha \phi)^\dagger \partial_t \epsilon_0 dx dt \right| \lesssim \sum_{k_1 > b_L} \sum_{k_2 = k_1 + O(1)} \left| \int_{I_1} \int_{\mathbb{R}^2} P_0(P_{k_1} \epsilon P_{k_2} (\partial_\alpha \phi^\dagger \partial^\alpha \phi))^\dagger \partial_t \epsilon_0 dx dt \right|.$$

In case of high-high interactions, we place $P_{k_2}(\partial_\alpha \phi^\dagger \partial^\alpha \phi)$ into $L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}$, while in case of low-high and high-low interactions, we put $P_{k_2}(\partial_\alpha \phi^\dagger \partial^\alpha \phi)$ into $L_t^{\frac{5}{3}} L_x^{\frac{5}{3}}$. Thus, in the former case we obtain from the null form estimate (4.1), the bootstrap assumptions (7.13) and the exponential decay (7.12) of the S_k norms of ϕ for $k > b_L$ the bound

$$\begin{aligned} &\lesssim \sum_{k_1 > b_L} \sum_{k_3 > k_1} \|P_{k_1} \epsilon\|_{L_t^3 L_x^\infty[I_1]} \|P_{k_3} \phi\|_{S_{k_3}[I_1]}^2 \|\partial_t \epsilon_0\|_{L_t^\infty L_x^3[I_1]} \\ &\lesssim \sum_{k_1 > b_L} \sum_{k_3 > k_1} 2^{-\frac{1}{3}k_1} \epsilon_0 \delta_1^2 2^{-2\sigma(k_3 - b_L)} C \delta_2 2^{-\sigma b_L} \\ &\ll C^2 \delta_2^2 2^{-2\sigma b_L} \end{aligned}$$

for $\delta_1 > 0$ sufficiently small. In the latter case, we similarly infer the estimate

$$\begin{aligned} &\lesssim \sum_{k_1 > b_L} \sum_{k_3 \leq k_1 + O(1)} \|P_{k_1} \epsilon\|_{L_t^{\frac{5}{2}} L_x^\infty[I_1]} 2^{\frac{1}{10}k_1} 2^{\frac{1}{10}k_3} \|P_{k_1} \phi\|_{S_{k_1}[I_1]} \|P_{k_3} \phi\|_{S_{k_3}[I_1]} \|\partial_t \epsilon_0\|_{L_t^\infty L_x^{\frac{5}{2}}[I_1]} \\ &\lesssim \sum_{k_1 > b_L} \sum_{k_3 \leq k_1 + O(1)} 2^{-\frac{1}{5}k_1} \epsilon_0 2^{\frac{1}{10}k_1} 2^{\frac{1}{10}k_3} \delta_1 2^{-\sigma(k_1 - b_L)} \|P_{k_3} \phi\|_{S_{k_3}[I_1]} C \delta_2 2^{-\sigma b_L} \\ &\lesssim \sum_{k_1 > b_L} 2^{-\frac{1}{5}k_1} \epsilon_0 \delta_1 2^{-\sigma(k_1 - b_L)} C(E_{crit}) C \delta_2 2^{-\sigma b_L} \\ &\ll C^2 \delta_2^2 2^{-2\sigma b_L}. \end{aligned}$$

Next, for the contribution of the second term in (7.20)

$$\left| \int_{I_1} \int_{\mathbb{R}^2} P_0(P_{[-10, b_L]} \epsilon \partial_\alpha \phi^\dagger \partial^\alpha \phi)^\dagger \partial_t \epsilon_0 dx dt \right|,$$

we argue similarly and use that

$$\|P_{k_1} \epsilon\|_{L_t^q L_x^\infty} \lesssim C \delta_2 2^{-\frac{1}{q}k_1 - \sigma(b_L - k_1)} \text{ for } k_1 < b_L$$

and a divisibility argument (from $\partial_\alpha \phi^\dagger \partial^\alpha \phi$) to force smallness. Finally, in order to treat the third term in (7.20), for each frequency $k \leq -10$ we consider the expression

$$\left| \int_{I_1} \int_{\mathbb{R}^2} P_0(P_k \epsilon P_0(\partial_\alpha \phi^\dagger \partial^\alpha \phi))^\dagger \partial_t \epsilon_0 dx dt \right|$$

and now further distinguish for the radial variable the cases $r \leq 1$ and $r > 1$.

When $r \leq 1$ in case of low-high and high-low interactions in $P_0(\partial_\alpha \phi^\dagger \partial^\alpha \phi)$ we easily estimate

$$\begin{aligned}
& \left| \int_{I_1} \int_{\mathbb{R}^2} \chi_{\{r \leq 1\}} P_0(P_k \in P_0(\partial_\alpha \phi^\dagger \partial^\alpha \phi))^\dagger \partial_t \epsilon_0 \, dx \, dt \right| \\
& \lesssim \sum_{\ell < 0} \left\| \chi_{\{r \sim 2^\ell\}} r^{+\frac{1}{2}} \right\|_{L_x^6} \|P_k \in\|_{L_t^\infty L_x^\infty[I_1]} \|P_0(\partial_\alpha \phi \partial^\alpha \phi)\|_{L_t^2 L_x^3[I_1]} \sup_{\ell \in \mathbb{Z}} \left\| \chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \partial_t \epsilon_0 \right\|_{L_t^2 L_x^2[I_1]} \\
& \lesssim \|P_k \in\|_{S_k[I_1]} \|P_0(\partial_\alpha \phi \partial^\alpha \phi)\|_{L_t^2 L_x^3[I_1]} \|\epsilon_0\|_{S_0[I_1]} \\
& \lesssim 2^{\sigma k} C^2 \delta_2^2 2^{-2\sigma b_L} \|P_0(\partial_\alpha \phi \partial^\alpha \phi)\|_{L_t^2 L_x^3[I_1]}.
\end{aligned}$$

Then one can sum over $k \leq -10$ and smallness is obtained from divisibility of the $\|P_0(\partial_\alpha \phi \partial^\alpha \phi)\|_{L_t^2 L_x^3[I_1]}$ norm. In case of high-high interactions in $P_0(\partial_\alpha \phi^\dagger \partial^\alpha \phi)$, we use Hölder's inequality to place $P_k \in$ into $L_t^\infty L_x^\infty[I_1]$, the null form $P_0(\partial_\alpha \phi^\dagger \partial^\alpha \phi)$ into $L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}[I_1]$ and $\partial_t \epsilon_0$ into $L_t^3 L_x^\infty[I_1]$, gaining smallness from divisibility of the $L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}[I_1]$ norm in this case.

We can therefore now turn to the more subtle case when $r > 1$. We first treat the high-high interactions in $P_0(\partial_\alpha \phi^\dagger \partial^\alpha \phi)$. Here we use Strauss' improved Sobolev embedding to estimate

$$\begin{aligned}
& \left| \int_{I_1} \int_{\mathbb{R}^2} \chi_{\{r > 1\}} P_0(P_k \in \sum_{k_1 > 0} P_0(\partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_1+O(1)}))^\dagger \partial_t \epsilon_0 \, dx \, dt \right| \\
& \lesssim \sum_{\ell > 0} \left\| \chi_{\{r \sim 2^\ell\}} P_k \in \right\|_{L_t^\infty L_x^\infty} \left\| \sum_{k_1 > 0} r^{+\frac{1}{2}} (\partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_1+O(1)}) \right\|_{L_t^2 L_x^2[I_1]} \sup_{\ell \in \mathbb{Z}} \left\| \chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \partial_t \epsilon_0 \right\|_{L_t^2 L_x^2[I_1]} \\
& \lesssim \sum_{\ell > 0} \min\{2^{-\frac{1}{2}(\ell+k)}, 1\} \|P_k \in\|_{S_k[I_1]} \|\epsilon_0\|_{S_0[I_1]} \left\| \sum_{k_1 > 0} r^{+\frac{1}{2}} (\partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_1+O(1)}) \right\|_{L_t^2 L_x^2[I_1]} \\
& \lesssim (1 + |k|) 2^{\sigma k} C^2 \delta_2^2 2^{-2\sigma b_L} \left\| \sum_{k_1 > 0} r^{+\frac{1}{2}} (\partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_1+O(1)}) \right\|_{L_t^2 L_x^2[I_1]}.
\end{aligned}$$

This bound can be summed over $k \leq -10$ and we obtain smallness since the $L_t^2 L_x^2[I_1]$ norm here has the divisibility property thanks to the weighted null form estimate (4.3). In case of low-high interactions in $P_0(\partial_\alpha \phi^\dagger \partial^\alpha \phi)$ (and then analogously for high-low interactions), we write

$$\begin{aligned}
& \int_{I_1} \int_{\mathbb{R}^2} \chi_{\{r > 1\}} P_0(P_k \in \sum_{k_1 < O(1)} P_0(\partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_0))^\dagger \partial_t \epsilon_0 \, dx \, dt \\
& = \int_{I_1} \int_{\mathbb{R}^2} \chi_{\{r > 1\}} P_0(P_k \in \sum_{k < k_1 < O(1)} P_0(\partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_0))^\dagger \partial_t \epsilon_0 \, dx \, dt \\
& \quad + \sum_{k_1 \leq k} \int_{I_1} \int_{\mathbb{R}^2} \chi_{\{r > 1\}} P_0(P_k \in P_0(\partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_0))^\dagger \partial_t \epsilon_0 \, dx \, dt.
\end{aligned}$$

Then we can bound the first term on the right-hand side exactly as in the high-high case only that we pay a price of $|k|^2$, which can still be absorbed by the gain $2^{\sigma k}$ and then summed over $k \leq -10$. For the second term on the right-hand side we distribute the weight differently, more precisely for some $0 < \lambda < \frac{1}{2}$ we again

use Strauss' improved Sobolev embedding and estimate by

$$\begin{aligned}
& \sum_{\ell > 0} \sum_{k_1 \leq k} \|\chi_{\{r \sim 2^\ell\}} r^{\frac{1}{2}-\lambda} P_k \epsilon\|_{L_t^\infty L_x^\infty[I_1]} \|r^{+\lambda} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_0\|_{L_t^2 L_x^2[I_1]} \sup_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \partial_t \epsilon_0\|_{L_t^2 L_x^2} \\
& \lesssim \sum_{k_1 \leq k} \sum_{\ell > 0} \min\{2^{-\lambda(\ell+k)}, 1\} 2^{-(\frac{1}{2}-\lambda)k} \|P_k \epsilon\|_{S_k[I_1]} \|r^{+\lambda} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_0\|_{L_t^2 L_x^2[I_1]} \|\epsilon_0\|_{S_0[I_1]} \\
& \lesssim \sum_{k_1 \leq k} (1 + |k|) 2^{-(\frac{1}{2}-\lambda)(k-k_1)} \|P_k \epsilon\|_{S_k[I_1]} 2^{-(\frac{1}{2}-\lambda)k_1} \|r^{+\lambda} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_0\|_{L_t^2 L_x^2[I_1]} \|\epsilon_0\|_{S_0[I_1]} \\
& \lesssim (1 + |k|) 2^{\sigma k} C^2 \delta_2^2 2^{-2\sigma b_L} \left(\sum_{k_1 \leq 0} 2^{-2(\frac{1}{2}-\lambda)k_1} \|r^{+\lambda} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_0\|_{L_t^2 L_x^2[I_1]}^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

The last line can then be summed over $k \leq -10$ and by the weighted null form estimate (4.3) the last factor here has the divisibility property yielding smallness.

(ii) *Contribution of $P_0(\phi \partial_\alpha \epsilon^\dagger \partial^\alpha \phi)$.* Here we split into

$$\begin{aligned}
\int_{I_1} \int_{\mathbb{R}^2} P_0(\phi \partial_\alpha \epsilon^\dagger \partial^\alpha \phi)^\dagger \partial_t \epsilon_0 \, dx \, dt &= \int_{I_1} \int_{\mathbb{R}^2} P_0(\phi_{\leq -10} \partial_\alpha \epsilon^\dagger \partial^\alpha \phi)^\dagger \partial_t \epsilon_0 \, dx \, dt \\
&\quad + \int_{I_1} \int_{\mathbb{R}^2} P_0(\phi_{> -10} \partial_\alpha \epsilon^\dagger \partial^\alpha \phi)^\dagger \partial_t \epsilon_0 \, dx \, dt.
\end{aligned}$$

Since the second term is truly quadrilinear and thus easier to treat, we only turn to the details of the estimate of the first term, which we may schematically write as

$$(7.21) \quad \int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi) \phi_{\leq -10}^\dagger \partial_t \epsilon_0 \, dx \, dt.$$

Since $\phi + \epsilon$ and ϕ are wave maps into the sphere, we have the following geometric identity at our disposal

$$0 = (\phi + \epsilon)^\dagger \partial_t (\phi + \epsilon) = \phi^\dagger \partial_t \epsilon + \epsilon^\dagger \partial_t \phi + \epsilon^\dagger \partial_t \epsilon.$$

Upon applying the projection P_0 and decomposing $P_0(\phi^\dagger \partial_t \epsilon)$ into a Littlewood-Paley trichotomy, we find that we can schematically rewrite the term (7.21) as

$$\int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi) \left(-\phi_0^\dagger \partial_t \epsilon_{\leq -10} - \sum_{k_1 > 0} P_0(\phi_{k_1}^\dagger \partial_t \epsilon_{k_1}) - P_0(\epsilon^\dagger \partial_t \phi) - P_0(\epsilon^\dagger \partial_t \epsilon) \right) dx \, dt.$$

This leads to four contributions, which we now estimate separately.

(ii.1) We begin with the term

$$\int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi) \phi_0^\dagger \partial_t \epsilon_{\leq -10} \, dx \, dt.$$

In case of high-low or low-high interactions in $P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi)$, we use the null form estimate (4.1) to place $P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi)$ into $L_t^{\frac{5}{3}} L_x^{\frac{5}{3}}$ and put ϕ_0 as well as $\partial_t \epsilon_{\leq -10}$ into $L_t^5 L_x^5$, gaining smallness from divisibility of $\|\phi_0\|_{L_t^5 L_x^5}$. In case of high-high interactions in $P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi)$ we integrate by parts to move a derivative ∂_α away from ϵ , which leads to the terms

$$\int_{I_1} \int_{\mathbb{R}^2} P_0(\epsilon^\dagger \square \phi) \phi_0^\dagger \partial_t \epsilon_{\leq -10} \, dx \, dt + \int_{I_1} \int_{\mathbb{R}^2} P_0(\epsilon^\dagger \partial^\alpha \phi) \partial_\alpha \phi_0^\dagger \partial_t \epsilon_{\leq -10} \, dx \, dt + \int_{I_1} \int_{\mathbb{R}^2} P_0(\epsilon^\dagger \partial^\alpha \phi) \phi_0^\dagger \partial_\alpha \partial_t \epsilon_{\leq -10} \, dx \, dt,$$

where in the first term we have to reinsert the wave maps equation for ϕ . Then one can essentially argue as in (i) to bound these terms.

(ii.2) Next, we consider the term

$$\int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi) \sum_{k_1 > 0} P_0(\phi_{k_1}^\dagger \partial_t \epsilon_{k_1}) \, dx \, dt.$$

In case of high-low interactions in $P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi)$ (and then analogously for low-high interactions), we split into

$$\sum_{0 < k_1 \leq b_L} \int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon_0^\dagger \partial^\alpha \phi_{\leq -10}) P_0(\phi_{k_1}^\dagger \partial_t \epsilon_{k_1}) dx dt + \sum_{k_1 > b_L} \int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon_0^\dagger \partial^\alpha \phi_{\leq -10}) P_0(\phi_{k_1}^\dagger \partial_t \epsilon_{k_1}) dx dt.$$

Then it is easy to see that in the first summand we can restrict to $r > 1$, which allows us to estimate via the weighted null form estimate (4.3) for some $0 < \lambda < \frac{1}{2}$,

$$\begin{aligned} & \left| \sum_{0 < k_1 \leq b_L} \int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon_0^\dagger \partial^\alpha \phi_{\leq -10}) \chi_{\{r > 1\}} P_0(\phi_{k_1}^\dagger \partial_t \epsilon_{k_1}) dx dt \right| \\ & \lesssim \|r^{+\lambda} P_0(\partial_\alpha \epsilon_0^\dagger \partial^\alpha \phi_{\leq -10})\|_{L_t^2 L_x^2[I_1]} \sum_{0 \leq k_1 < b_L} \|\chi_{\{r > 2^{-k_1}\}} r^{-\lambda} \phi_{k_1}\|_{L_t^2 L_x^\infty[I_1]} \|\partial_t \epsilon_{k_1}\|_{L_t^\infty L_x^2[I_1]} \\ & \lesssim \|\epsilon_0\|_{S_0[I_1]} \|\phi\|_{S[I_1]} \left(\sum_{0 < k_1 \leq b_L} 2^{2(\frac{1}{2}-\lambda)k_1} \|\chi_{\{r > 2^{-k_1}\}} r^{-\lambda} \phi_{k_1}\|_{L_t^2 L_x^\infty[I_1]}^2 \right)^{\frac{1}{2}} \left(\sum_{0 < k_1 \leq b_L} 2^{-2(\frac{1}{2}-\lambda)k_1} C^2 \delta_2^2 2^{-2\sigma(b_L-k_1)} \right)^{\frac{1}{2}} \\ & \lesssim C^2 \delta_2^2 2^{-2\sigma b_L} \left(\sum_{0 < k_1 \leq b_L} 2^{2(\frac{1}{2}-\lambda)k_1} \|\chi_{\{r > 2^{-k_1}\}} r^{-\lambda} \phi_{k_1}\|_{L_t^2 L_x^\infty[I_1]}^2 \right)^{\frac{1}{2}} \end{aligned}$$

and here the last factor has the divisibility property yielding smallness. For the second summand we use Bernstein and the null form estimate (4.1) to bound by

$$\begin{aligned} & \left| \sum_{k_1 > b_L} \int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon_0^\dagger \partial^\alpha \phi_{\leq -10}) P_0(\phi_{k_1}^\dagger \partial_t \epsilon_{k_1}) dx dt \right| \\ & \lesssim \|\partial_\alpha \epsilon_0^\dagger \partial^\alpha \phi_{\leq -10}\|_{L_t^{\frac{5}{3}} L_x^{\frac{5}{3}}[I_1]} \sum_{k_1 > b_L} \|\phi_{k_1}\|_{L_t^{\frac{5}{2}} L_x^\infty[I_1]} \|\partial_t \epsilon_{k_1}\|_{L_t^\infty L_x^2[I_1]} \\ & \lesssim \|\epsilon_0\|_{S_0[I_1]} \|\phi\|_{S[I_1]} \sum_{k_1 > b_L} 2^{-\frac{5}{2}k_1} \delta_1 2^{-\sigma(k_1-b_L)} \|\epsilon_{k_1}\|_{S_{k_1}[I_1]} \\ & \lesssim C \delta_2 2^{-\sigma b_L} C(E_{crit}) \delta_1 2^{-\frac{5}{2}b_L} \epsilon_0^{\frac{1}{2}} \\ & \ll C^2 \delta_2^2 2^{-2\sigma b_L}. \end{aligned}$$

In case of high-high interactions in $P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi)$ we have to integrate by parts to move a derivative ∂_α away from ϵ and then one can argue essentially as in (i).

(ii.3) Here we further decompose the term $P_0(\epsilon^\dagger \partial_t \phi)$ into a Littlewood-Paley trichotomy so that we schematically have

$$\begin{aligned} \int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi) P_0(\epsilon^\dagger \partial_t \phi) dx dt &= \int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi) P_0(\epsilon_{\leq -10}^\dagger \partial_t \phi_0) dx dt \\ &+ \int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi) P_0(\epsilon_0^\dagger \partial_t \phi_{\leq -10}) dx dt \\ &+ \sum_{k_1 > 0} \int_{I_1} \int_{\mathbb{R}^2} P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi) P_0(\epsilon_{k_1}^\dagger \partial_t \phi_{k_1}) dx dt. \end{aligned} \tag{7.22}$$

For the first term on the right-hand side of (7.22) for high-low interactions in $P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi)$ (and similarly for low-high interactions) we can proceed as we did for the third term in (7.20), where now $\partial_\alpha \epsilon^\dagger \partial^\alpha \phi$ replaces $\partial_\alpha \phi^\dagger \partial^\alpha \phi$. Correspondingly, the divisibility part to achieve smallness has to be modified a bit, and in fact has to come from the last factor $\partial_t \phi_0$ which we place in the local energy decay space. Specifically, we arrive at

the expression

$$\left(\sum_{\ell \in \mathbb{Z}} (1 + \ell^2)^{-1} \left\| \chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \partial_t \phi_0 \right\|_{L_t^2 L_x^2[I_1]}^2 \right)^{\frac{1}{2}},$$

which has the divisibility property here. For high-high interactions in $P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi)$ one again has to integrate by parts to move a derivative ∂_α away from ϵ .

The second term on the right-hand side of (7.22) can be treated easily using the null form estimate (4.1) and Strichartz estimates in case of high-low and low-high interactions in $P_0(\partial_\alpha \epsilon^\dagger \partial^\alpha \phi)$, while for high-high interactions one has to integrate by parts again.

Finally, the third term on the right-hand side of (7.22) can be dealt with analogously to (ii.2).

(ii.4) In this case we can essentially proceed as in (ii.3) only that it is easier to achieve smallness thanks to the extra factor ϵ .

(iii) *Contribution of $P_0(\epsilon \partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon)$.* As usual we split into

$$\begin{aligned} \int_{I_1} \int_{\mathbb{R}^2} P_0(\epsilon \partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon)^\dagger \partial_t \epsilon_0 \, dx \, dt &= \int_{I_1} \int_{\mathbb{R}^2} P_0(\epsilon_{\leq -10} \partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon)^\dagger \partial_t \epsilon_0 \, dx \, dt \\ &+ \sum_{k_1 > -10} \sum_{k_2 = k_1 + O(1)} \int_{I_1} \int_{\mathbb{R}^2} P_0(P_{k_1} \epsilon P_{k_2} (\partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon))^\dagger \partial_t \epsilon_0 \, dx \, dt. \end{aligned}$$

For the first term on the right-hand side we may proceed as at the end of (i) to infer the desired bound. For the second term we distinguish high-high interactions in $P_{k_2}(\partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon)$ from high-low and low-high interactions.

In the former case we place $P_{k_2}(\partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon)$ into $L_t^{\frac{3}{2}} L_x^{\frac{3}{2}}$ and use the null form estimate (4.1) to infer

$$\begin{aligned} &\left| \sum_{k_1 > -10} \sum_{k_2 = k_1 + O(1)} \int_{I_1} \int_{\mathbb{R}^2} P_0(P_{k_1} \epsilon P_{k_2} (\partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon))^\dagger \partial_t \epsilon_0 \, dx \, dt \right| \\ &\lesssim \sum_{k_3 > -10} \|P_{k_3} \epsilon\|_{S_{k_1}[I_1]}^2 \sum_{k_1 > -10} \|P_{k_1} \epsilon\|_{L_t^3 L_x^\infty[I_1]} \|\partial_t \epsilon_0\|_{L_t^\infty L_x^3[I_1]} \\ &\lesssim \varepsilon_0 \sum_{k_1 > -10} 2^{-\frac{1}{3}k_1} \|P_{k_1} \epsilon\|_{S_{k_1}[I_1]} \|\partial_t \epsilon_0\|_{S_0[I_1]}. \end{aligned}$$

By further distinguishing the frequency regimes and exploiting the smallness and decay properties (7.14)–(7.16) of $\{c_k\}_{k \in \mathbb{Z}}$ and $\{d_k\}_{k \in \mathbb{Z}}$, we find that

$$\begin{aligned} &\sum_{k_1 > -10} 2^{-\frac{1}{3}k_1} \|P_{k_1} \epsilon\|_{S_{k_1}[I_1]} \|\partial_t \epsilon_0\|_{S_0[I_1]} \\ &\lesssim \sum_{-10 < k_1 \leq b_L} 2^{-(\frac{1}{3}-\sigma)k_1} C^2 \delta_2^2 2^{-2\sigma b_L} + \sum_{b_L < k_1 \leq b_L + M} 2^{-\frac{1}{3}k_1} c_{k_1} C \delta_2 2^{-\sigma b_L} + \sum_{k_1 > b_L + M} 2^{-\frac{1}{3}k_1} c_{k_1} C \delta_2 2^{-\sigma b_L} \\ &\lesssim C^2 \delta_2^2 2^{-2\sigma b_L} + \left(\sum_{b_L < k_1 \leq b_L + M} c_{k_1}^2 \right)^{\frac{1}{2}} \delta_2 2^{-(\sigma + \frac{1}{3})b_L} + 2^{-\frac{1}{3}M} \varepsilon_0 C \delta_2 2^{-(\sigma + \frac{1}{3})b_L} \\ &\lesssim C^2 \delta_2^2 2^{-2\sigma b_L}. \end{aligned}$$

In total, for sufficiently small $\varepsilon_0 > 0$ we obtain the desired bound

$$\left| \sum_{k_1 > -10} \sum_{k_2 = k_1 + O(1)} \int_{I_1} \int_{\mathbb{R}^2} P_0(P_{k_1} \epsilon P_{k_2} (\partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon))^\dagger \partial_t \epsilon_0 \, dx \, dt \right| \lesssim \varepsilon_0 C^2 \delta_2^2 2^{-2\sigma b_L} \ll C^2 \delta_2^2 2^{-2\sigma b_L}.$$

In the case of high-low or low-high interactions, we place $P_{k_2}(\partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon)$ into $L_t^{\frac{3}{2}+} L_x^{\frac{3}{2}+}$ and then proceed similarly as above.

(iv) *Contributions of $P_0(\epsilon \partial_\alpha \epsilon^\dagger \partial^\alpha \phi)$ and $P_0(\phi \partial_\alpha \epsilon^\dagger \partial^\alpha \epsilon)$* : These terms do not offer anything new and can be treated similarly as above, which finishes the proof of Lemma 7.10. \square

Next, we recover the full envelope bound. We may assume that $k = 0$ and now want to infer the improved bound $\|\epsilon_0\|_{S_0[I_1]} \leq \frac{C}{2}(c_0 + d_0)$. To this end we will use the “renormalized difference equation” for ϵ_0 given by

$$\square(U_{\leq -10}^{(\phi+\epsilon)}(\phi + \epsilon)_0) - \square(U_{\leq -10}^{(\phi)}\bar{\phi}_0) \equiv F_0^{(\phi+\epsilon)} - F_0^{(\epsilon)},$$

where $U_{\leq h}^{(\phi+\epsilon)}$ and $U_{\leq h}^{(\phi)}$ denote the gauge transformations defined by $\phi + \epsilon$, respectively ϕ . We will proceed in several steps which we briefly sketch before we turn to the details.

(a) First we show that for some $D_1 \gg 1$ with $1 \ll D_0 \ll D_1 \ll C$, we have for any $0 \leq \alpha < 1$ that

$$(7.23) \quad \sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k_0|} \|P_{k_0} \square(U_{\leq -10}^{(\phi+\epsilon)}(\phi + \epsilon)_0 - U_{\leq -10}^{(\phi)}\bar{\phi}_0)\|_{L_t^1 L_x^2[I_1]} \leq D_1(c_0 + d_0).$$

Note that here it is again necessary to control the above weighted sum over all frequency outputs since the frequency localizations of $(\phi + \epsilon)_0$ and $\bar{\phi}_0$ to frequency ~ 1 are only approximately preserved by the gauge transformations up to exponentially decaying tails. By the energy estimate (3.1) we can then conclude that

$$(7.24) \quad \sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k_0|} \|P_{k_0}(U_{\leq -10}^{(\phi+\epsilon)}(\phi + \epsilon)_0 - U_{\leq -10}^{(\phi)}\bar{\phi}_0)\|_{S_{k_0}[I_1]} \lesssim D_1(c_0 + d_0).$$

(b) Next we show that we have

$$(7.25) \quad \sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k_0|} \|P_{k_0}(U_{\leq -10}^{(\phi+\epsilon)} - U_{\leq -10}^{(\phi)})\bar{\phi}_0\|_{S_{k_0}[I_1]} \lesssim D_1(c_0 + d_0).$$

This step crucially relies on the improved bounds on the kinetic energy of ϵ_k for the low frequencies $k < b_L$, which we have already established in Lemma 7.10.

(c) Combining (7.24) and (7.25) we immediately obtain

$$(7.26) \quad \sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k_0|} \|P_{k_0} U_{\leq -10}^{(\phi+\epsilon)}((\phi + \epsilon)_0 - \bar{\phi}_0)\|_{S_{k_0}[I_1]} \lesssim D_1(c_0 + d_0).$$

(d) Then we write

$$(\phi + \epsilon)_0 - \bar{\phi}_0 = (U_{\leq -10}^{(\phi+\epsilon)})^\dagger U_{\leq -10}^{(\phi+\epsilon)}((\phi + \epsilon)_0 - \bar{\phi}_0)$$

and decompose into a Littlewood-Paley trichotomy. Using the bound (7.26) we may then proceed as in the proof of Proposition 6.3 to infer that for some $D_2 \gg 1$ with $1 \ll D_0 \ll D_1 \ll D_2 \ll C$,

$$(7.27) \quad \|(\phi + \epsilon)_0 - \bar{\phi}_0\|_{S_0[I_1]} \leq D_2(c_0 + d_0).$$

Here we do not have to work with the modified version of the gauge transformations as introduced in Proposition 6.3 because we already have that $\|\phi\|_{S[I_1]} \leq C(E_{crit})$. We will correspondingly not provide further details on this step.

(e) Finally, we invoke the bound (7.27) and exploit the inherently multilinear structure of

$$(\phi + \epsilon)_0 - \bar{\phi}_0 - \epsilon_0$$

to conclude the desired improved bound

$$\|\epsilon_0\|_{S_0[I_1]} \leq \frac{C}{2}(c_0 + d_0).$$

We now provide the details of the key steps (a), (b) and (e) in separate lemmas, starting with (a).

Lemma 7.11. *If ϵ satisfies the bootstrap assumption (7.13) for some $C \equiv C(E_{crit}) \gg 1$ sufficiently large, then we have for some $1 \ll D_1 \ll C$ that for all $0 \leq \alpha < 1$,*

$$(7.28) \quad \sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k_0|} \|P_{k_0} \square(U_{\leq -10}^{(\phi+\epsilon)}(\phi + \epsilon)_0 - U_{\leq -10}^{(\phi)}\bar{\phi}_0)\|_{L_t^1 L_x^2[I_1]} \leq D_1(c_0 + d_0).$$

Proof. We begin by observing that it suffices to prove

$$(7.29) \quad \|\square(U_{\leq -10}^{(\phi+\epsilon)}(\phi + \epsilon)_0 - U_{\leq -10}^{(\phi)}\bar{\phi}_0)\|_{L_t^1 L_x^2[I_1]} \leq D_1(c_0 + d_0).$$

Then the bound on the weighted sum (7.28) over all output frequencies can be established as in the proof of Proposition 6.3.

For each characteristic component of the renormalized wave maps nonlinearity as summarized in Proposition 5.1, we now have to establish the estimate (7.29) for the difference of the corresponding expression for $\phi + \epsilon$ and ϕ . Here we show this in detail for the first term on the right-hand side of (5.32), namely when a derivative falls on the lowest frequency; for all other terms one has to proceed similarly. We are therefore now facing a difference term of the schematic form

$$U_{\leq -10}^{(\phi+\epsilon)} \sum_{k_2 \leq -10} (\phi + \epsilon)_{k_2-10 < \cdot \leq -10} \partial_\alpha (\phi + \epsilon)_{k_2}^\dagger \partial^\alpha (\phi + \epsilon)_0 - U_{\leq -10}^{(\phi)} \sum_{k_2 \leq -10} \phi_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0.$$

This leads to a number of expressions which we are now estimating separately.

(i) *Contribution of $(U_{\leq -10}^{(\phi+\epsilon)} - U_{\leq -10}^{(\phi)}) \sum_{k_2 \leq -10} \phi_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0$.* We distinguish further between the relation of the frequency scale 0 to b_L , which is the upper endpoint of the “essential” frequency support of ϕ .

(i.1) $0 \geq b_L$. This case is straightforward on account of the smallness $\|P_k \phi\|_{S_k} \lesssim \delta_1 2^{-\sigma(k-b_L)}$ for $k > b_L$. We simply use the exact orthogonality of the gauge transformations and the trilinear estimate (4.4) to bound

$$\begin{aligned} \|(U_{\leq -10}^{(\phi+\epsilon)} - U_{\leq -10}^{(\phi)}) \sum_{k_2 \leq -10} \phi_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0\|_{L_t^1 L_x^2[I_1]} &\lesssim \sum_{k_2 \leq -10} \|\phi_{k_2}\|_{S_{k_2}[I_1]}^2 \|\phi_0\|_{S_0[I_1]} \\ &\lesssim C(E_{crit}) \delta_1 2^{+\sigma b_L} \\ &\ll \delta_2 2^{+\sigma b_L}. \end{aligned}$$

(i.2) $0 < b_L$. Here we have to argue more carefully, since we now have to gain smallness from the difference of the gauge transformations $U_{\leq -10}^{(\phi+\epsilon)} - U_{\leq -10}^{(\phi)}$. By the defining ODE (5.28) for the gauge transformations we have that

$$U_{\leq -10}^{(\phi+\epsilon)} - U_{\leq -10}^{(\phi)} = \int_{-\infty}^{-10} U_{\leq h}^{(\phi+\epsilon)} (B_h^{(\phi+\epsilon)} - B_h^{(\phi)}) dh + \int_{-\infty}^{-10} (U_{\leq h}^{(\phi+\epsilon)} - U_{\leq h}^{(\phi)}) B_h^{(\phi)} dh.$$

By further expanding the second term on the right-hand side, we obtain an expansion containing only terms involving differences $B_h^{(\phi+\epsilon)} - B_h^{(\phi)}$, namely

$$(7.30) \quad U_{\leq -10}^{(\phi+\epsilon)} - U_{\leq -10}^{(\phi)} = \sum_{n=1}^{\infty} \int_{-\infty}^{-10} \int_{-\infty}^{h_1} \cdots \int_{-\infty}^{h_{n-1}} U_{\leq h_n}^{(\phi+\epsilon)} (B_{h_n}^{(\phi+\epsilon)} - B_{h_n}^{(\phi)}) B_{h_{n-1}}^{(\phi)} \cdots B_{h_1}^{(\phi)} dh_n \cdots dh_1.$$

Since the repeated expansion leads to difference terms of lower and lower frequencies, this infinite sum is convergent due to the assumed exponential decay of d_k for $k < b_L$. In view of the definition (5.27) of $B_h^{(\phi+\epsilon)}$ and $B_h^{(\phi)}$, we may therefore replace (7.30) by the schematic expression

$$\sum_{k_1 \leq -10} L(\phi_{\leq k_1-10}, \epsilon_{k_1}) + L(\epsilon_{\leq k_1-10}, \phi_{k_1}) + L(\epsilon_{\leq k_1-10}, \epsilon_{k_1}).$$

Noting that by a simple divisibility argument we may assume that

$$\left\| \sum_{k_2 \leq -10} \phi_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0 \right\|_{L_t^1 L_x^2[I_1]} \ll 1,$$

we find that

$$\begin{aligned}
& \left\| (U_{\leq -10}^{(\phi+\epsilon)} - U_{\leq -10}^{(\phi)}) \sum_{k_2 \leq -10} \phi_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0 \right\|_{L_t^1 L_x^2[I_1]} \\
& \lesssim \left\| \left(\sum_{k_1 \leq -10} L(\phi_{\leq k_1-10}, \epsilon_{k_1}) + L(\epsilon_{\leq k_1-10}, \phi_{k_1}) + L(\epsilon_{\leq k_1-10}, \epsilon_{k_1}) \right) \sum_{k_2 \leq -10} \phi_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0 \right\|_{L_t^1 L_x^2[I_1]} \\
& \lesssim \sum_{k_1 \leq -10} \|\epsilon_{k_1}\|_{L_t^\infty L_x^\infty[I_1]} \left\| \sum_{k_2 \leq -10} \phi_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0 \right\|_{L_t^1 L_x^2[I_1]} \\
& \lesssim \sum_{k_1 \leq -10} \delta_2 2^{-\sigma(b_L - k_1)} \left\| \sum_{k_2 \leq -10} \phi_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0 \right\|_{L_t^1 L_x^2[I_1]} \\
& \ll \delta_2 2^{-\sigma b_L}.
\end{aligned}$$

(ii) *Contribution of $U_{\leq -10}^{(\phi+\epsilon)} \sum_{k_2 \leq -10} \epsilon_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0$.* Again we further distinguish between the relation of the frequency scale 0 and b_L .

(ii.1) $0 \geq b_L$. Here we immediately obtain from the smallness $\|\phi_k\|_{S_k} \lesssim \delta_1 2^{-\sigma(k-b_L)}$ for $k > b_L$ and the trilinear estimate (4.4) the desired bound

$$\begin{aligned}
\left\| U_{\leq -10}^{(\phi+\epsilon)} \sum_{k_2 \leq -10} \epsilon_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0 \right\|_{L_t^1 L_x^2[I_1]} & \lesssim \left(\sum_{k_1} \|\epsilon_{k_1}\|_{S_{k_1}[I_1]}^2 \right)^{\frac{1}{2}} \left(\sum_{k_2} \|\phi_{k_2}\|_{S_{k_2}[I_1]}^2 \right)^{\frac{1}{2}} \|\phi_0\|_{S_0[I_1]} \\
& \lesssim \varepsilon_0^{\frac{1}{2}} C(E_{crit}) \delta_1 2^{+\sigma b_L} \\
& \ll \delta_2 2^{+\sigma b_L}.
\end{aligned}$$

(ii.2) $0 < b_L$. By the trilinear estimate (4.4) here we have

$$\begin{aligned}
\left\| U_{\leq -10}^{(\phi+\epsilon)} \sum_{k_2 \leq -10} \epsilon_{k_2-10 < \cdot \leq -10} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_0 \right\|_{L_t^1 L_x^2[I_1]} & \lesssim \sum_{k_2 \leq -10} \sum_{k_2-10 < k_1 \leq -10} 2^{-\frac{1}{4}(k_1-k_2)} \|\epsilon_{k_1}\|_{S_{k_1}[I_1]} \|\phi_{k_2}\|_{S_{k_2}[I_1]} \|\phi_0\|_{S_0[I_1]} \\
& \lesssim \delta_2 2^{-\sigma b_L} C(E_{crit})
\end{aligned}$$

and then one forces smallness via a divisibility argument.

(iii) *Contribution of $U_{\leq -10}^{(\phi+\epsilon)} \sum_{k_2 \leq -10} \epsilon_{k_2-10 < \cdot \leq -10} \partial_\alpha \epsilon_{k_2}^\dagger \partial^\alpha \epsilon_0$.* In this case the smallness comes from the bound $\sum_{k \in \mathbb{Z}} c_k^2 \lesssim \varepsilon_0$. Indeed, by the trilinear estimate we obtain

$$\begin{aligned}
\left\| U_{\leq -10}^{(\phi+\epsilon)} \sum_{k_2 \leq -10} \epsilon_{k_2-10 < \cdot \leq -10} \partial_\alpha \epsilon_{k_2}^\dagger \partial^\alpha \epsilon_0 \right\|_{L_t^1 L_x^2[I_1]} & \lesssim \left(\sum_{k_2 \leq -10} \|\epsilon_{k_2}\|_{S_{k_2}[I_1]}^2 \right) \|\epsilon_0\|_{S_0[I_1]} \\
& \lesssim \varepsilon_0 C(c_0 + d_0) \\
& \leq D_1(c_0 + d_0).
\end{aligned}$$

(iv) *All other contributions can be estimated similarly to the above cases.* □

We proceed with step (b).

Lemma 7.12. *If ϵ satisfies the bootstrap assumption (7.13) for some $C \equiv C(E_{crit}) \gg 1$ sufficiently large, then we have for some $1 \ll D_0 \ll D_1 \ll C$ that for all $0 \leq \alpha < 1$*

$$(7.31) \quad \sum_{k_0 \in \mathbb{Z}} 2^{\alpha|k_0|} \|P_{k_0}((U_{\leq -10}^{(\phi+\epsilon)} - U_{\leq -10}^{(\phi)})\bar{\phi}_0)\|_{S_{k_0}[I_1]} \lesssim D_1(c_0 + d_0).$$

Proof. We have to prove the bound (7.31) separately for each component of our S norm. To this end we distinguish again between the relationship of the frequency scale 0 to b_L . In case that $b_L < 0$, we exploit the smallness $\|\phi_k\|_{S_k} \lesssim \delta_1 2^{-\sigma(k-b_L)}$ for $k > b_L$. Instead when $b_L \geq 0$, we have to invoke the already improved estimate (7.18) on the kinetic energy of ϵ_k for the low frequencies $k < b_L$. Moreover, we achieve control over the weighted sum over all frequencies $k_0 \in \mathbb{Z}$ in (7.31) by essentially playing out the frequencies using Bernstein's inequality. \square

Finally, we turn to step (e). Having established that $\|(\phi + \epsilon)_0 - \bar{\phi}_0\|_{S_0[I_1]} \leq D_2(c_0 + d_0)$ for some $D_2 \ll C$, we may now quickly infer the desired improved bound

$$\|\epsilon_0\|_{S_0[I_1]} \leq \frac{C}{2}(c_0 + d_0)$$

and thus close our bootstrap argument by exploiting the inherently multilinear structure of $(\phi + \epsilon)_0 - \bar{\phi}_0 - \epsilon_0$. Indeed, recalling the definition (5.25) of $\bar{\phi}$, we see that each multilinear expression in $(\phi + \epsilon)_0 - \bar{\phi}_0 - \epsilon_0$ contains at least one factor of ϵ . Then we again distinguish the relationship between the frequency scale 0 and b_L and either exploit the smallness $\|\phi_k\|_{S_k} \lesssim \delta_1 2^{-\sigma(k-b_L)}$ for $k \geq b_L$ or the already improved bound on the kinetic energy of ϵ_k for $k < b_L$, which completes the final step (e).

At this point we have thus achieved that for $C \equiv C(E_{crit}) \gg 1$ sufficiently large, we may infer by bootstrap on the time interval I_1 the improved bound

$$(7.32) \quad \|P_k \epsilon\|_{S_k[I_1]} \leq \frac{C}{2}(c_k + d_k) \quad \text{for } k \in \mathbb{Z}.$$

We would now like to repeat the above bootstrap argument on the remaining time intervals I_2, \dots, I_N in order to conclude that ϵ exists globally in time and satisfies uniform in n global-in-time S norm bounds. As alluded to before, here we encounter the danger that the energy of ϵ could keep growing as we move to later time intervals. However, by (7.32) the frequency profile of ϵ is essentially preserved along the evolution on the time interval I_1 . For this reason there cannot actually be much energy transfer between ϵ and ϕ , since ϕ is exponentially decaying for frequencies $k > b_L$. The next lemma indeed shows that the energy of ϵ is approximately preserved along the evolution on the interval I_1 . More precisely, the energy transfer is controlled by the smallness parameter δ_2 , which can be chosen arbitrarily small.

Lemma 7.13. *Assuming the bounds (7.32) on the evolution of ϵ on I_1 , it holds that*

$$(7.33) \quad \|\nabla_{t,x} \epsilon(t)\|_{L_x^2}^2 \leq \|\nabla_{t,x} \epsilon(0)\|_{L_x^2}^2 + C(E_{crit}) \delta_2 \quad \text{for } t \in I_1.$$

Proof. By energy conservation for the wave maps evolutions $\phi + \epsilon$ and ϕ on I_1 , we have that

$$\sum_{\alpha=0}^2 \|\partial_\alpha(\phi + \epsilon)(t)\|_{L_x^2}^2 = \text{const.} \quad \text{for } t \in I_1$$

and

$$\sum_{\alpha=0}^2 \|\partial_\alpha \phi(t)\|_{L_x^2}^2 = \text{const.} \quad \text{for } t \in I_1.$$

Moreover, since it holds that

$$\sum_{\alpha=0}^2 \|\partial_\alpha \epsilon(t)\|_{L_x^2}^2 = \sum_{\alpha=0}^2 \|\partial_\alpha(\phi + \epsilon)(t)\|_{L_x^2}^2 - \sum_{\alpha=0}^2 \|\partial_\alpha \phi(t)\|_{L_x^2}^2 - 2 \sum_{\alpha=0}^2 \int_{\mathbb{R}^2} (\partial_\alpha \epsilon^\dagger \partial_\alpha \phi)(t, x) dx,$$

it suffices to estimate on the time interval I_1 ,

$$\sum_{\alpha=0}^2 \int_{\mathbb{R}^2} \partial_\alpha \epsilon^\dagger \partial_\alpha \phi dx = \sum_{k \in \mathbb{Z}} \sum_{\alpha=0}^2 \int_{\mathbb{R}^2} P_k \partial_\alpha \epsilon^\dagger \tilde{P}_k \partial_\alpha \phi dx,$$

where \tilde{P}_k is a slightly fattened Littlewood-Paley projection such that $\tilde{P}_k P_k = P_k$. Using the bounds (7.32) on ϵ as well as the exponential decay (7.12) of ϕ for frequencies $k \geq b_L$, we therefore obtain on I_1 that

$$\begin{aligned} \left| \sum_{\alpha=0}^2 \int_{\mathbb{R}^2} \partial_\alpha \epsilon^\dagger \partial_\alpha \phi \, dx \right| &\lesssim \sum_{k \in \mathbb{Z}} \|P_k \nabla_{t,x} \epsilon\|_{L_t^\infty L_x^2[I_1]} \|\tilde{P}_k \nabla_{t,x} \phi\|_{L_t^\infty L_x^2[I_1]} \\ &\lesssim \sum_{k \in (-\infty, b_L) \cap \mathbb{Z}} \delta_2 2^{-\sigma(b_L - k)} \|\tilde{P}_k \nabla_{t,x} \phi\|_{L_t^\infty L_x^2[I_1]} + \sum_{k \in [b_L, +\infty) \cap \mathbb{Z}} (c_k + d_k) \delta_1 2^{-\sigma(k - b_L)} \\ &\lesssim_{E_{crit}} \delta_2 + \delta_1 \\ &\lesssim_{E_{crit}} \delta_2. \end{aligned}$$

□

Since the number of time intervals $N \equiv N(C_2, E_{crit})$ is controlled by the size of C_2 and E_{crit} , we can a priori ensure to choose the underlying Besov error δ_0 so small that the energy of ϵ never leaves the perturbative regime, i.e. it stays less than, say, $2\epsilon_0$. Hence, we can carry out the above bootstrap argument to infer S norm bounds on ϵ on all remaining time intervals I_2, \dots, I_N . This finishes the proof of Proposition 7.9. □

7.4. Selecting concentration profiles and adding the first large frequency atom.

7.4.1. *Setting up a profile decomposition.* Having established control over the global evolution of the lowest frequency non-atomic part $\Pi_{\leq \mu_n^1 - \log(R_n)} \phi^n[0]$ in the previous subsection, we now “add in” the first large frequency atom $\phi^{n1}[0] = P_{[\mu_n^1 - \log(R_n), \mu_n^1 + \log(R_n)]} \phi^n[0]$ in the sense that we now attempt to evolve the data

$$\Pi_{\leq \mu_n^1 + \log(R_n)} \phi^n[0].$$

In this subsection we shall slightly abuse notation and write for simplicity

$$\phi^n[0] \equiv \Pi_{\leq \mu_n^1 + \log(R_n)} \phi^n[0].$$

Moreover, we denote by

$$u^n[0] \equiv \Pi_{\leq \mu_n^1 - \log(R_n)} \phi^n[0]$$

the data of the lowest frequency non-atomic part and by

$$\mu_n \equiv \mu_n^1 - \log(R_n)$$

the upper frequency cut-off delimiting the lowest frequency non-atomic part. Finally, we shall write

$$\epsilon^n[0] \equiv \Pi_{\leq \mu_n^1 + \log(R_n)} \phi^n[0] - \Pi_{\leq \mu_n^1 - \log(R_n)} \phi^n[0].$$

By rescaling we may assume that $\mu_n^1 \equiv 0$.

In fact, we shall first attempt to evolve a sequence of slightly modified data

$$\phi^{n, < K_*}[0] \equiv \Pi_{< K_*} \phi^n[0],$$

which have better frequency localization properties, and thence infer bounds on the evolutions of the original data $\Pi_{\leq \mu_n^1 + \log(R_n)} \phi^n[0]$. Correspondingly, we also introduce the notation

$$\epsilon^{n, < K_*}[0] \equiv \phi^{n, < K_*}[0] - u^n[0].$$

In order to obtain these bounds, we shall have to take advantage of a suitable profile decomposition, which however is quite delicate to achieve on account of the poor perturbative properties of the wave maps equation expressed in terms of the coordinate functions.

The very first step to pick the concentration profiles is to correctly identify their asymptotic behavior. To guess the correct ansatz, we shall throughout assume that the evolutions ϕ^n of the data $\phi^n[0]$ exist on some large time interval I . In particular, this means that the gauge transformations $U_{< k}^{(\phi^n)}$ associated with the evolutions ϕ^n are a priori well-defined. At this point it is important to observe that in the trilinear

estimate (4.4) we do not gain exponentially in the difference of the largest to the smallest frequency present. For this reason the low frequencies will exert a non-negligible influence on the high frequencies for large times, which we have to take into account by the correct choice of the concentration profiles. At frequency $k = 0$ we arrive at the equation

$$\square(U_{<0}^{(\phi^n)} \overline{\phi_0^n}) = \square(U_{<0}^{(\phi^n)} (u^n + \epsilon^n)_0) =: 2U_{<0}^{(\phi^n)} \mathcal{A}_{\alpha, low}(u^n) \partial^\alpha \epsilon_0^n + U_{<0}^{(\phi^n)} F_0(u^n, \epsilon^n),$$

where $\mathcal{A}_{\alpha, low}(u^n)$ incorporates all (perturbative) low frequency terms stemming purely from u^n for which there are no exponential frequency gains in the corresponding interactions in the wave maps nonlinearity. In view of the structure of the wave maps nonlinearity as detailed in Proposition 5.1, we arrive at the expression

$$\mathcal{A}_{\alpha, low}(u^n) = - \sum_{k_2 < \mu_n} \sum_{k_2 - 10 < k_1 < \mu_n} (u_{k_1}^n)(\partial_\alpha u_{k_2}^n)^\dagger - (\partial_\alpha u_{k_2}^n)(u_{k_1}^n)^\dagger + (U_{<\mu_n}^{(u^n)})^\dagger \partial_\alpha U_{<\mu_n}^{(u^n)} + A_{\alpha; <\mu_n}^{(u^n)},$$

where $U_{<\mu_n}^{(u^n)}$ is the gauge transformation defined by u^n and $A_{\alpha; <\mu_n}^{(u^n)}$ is the connection form (5.30) defined by u^n . Then we may write

$$U_{<0}^{(\phi^n)} \mathcal{A}_{\alpha, low}(u^n) \partial^\alpha \epsilon_0^n = U_{<0}^{(\phi^n)} \mathcal{A}_{\alpha, low}(u^n) (U_{<0}^{(\phi^n)})^\dagger \partial^\alpha (U_{<0}^{(\phi^n)} \epsilon_0^n) + error,$$

where we expect *error* to be a better term with a derivative on a low frequency term. Furthermore, we expect the high-frequency contribution to $U_{<0}^{(\phi^n)}$ coming from ϵ^n to be negligible, and so we replace the first term on the right by $U_{<0}^{(u^n)} \mathcal{A}_{\alpha, low}(u^n) (U_{<0}^{(u^n)})^\dagger \partial^\alpha (U_{<0}^{(\phi^n)} \epsilon_0^n)$. In fact, due to the rapid decay of u^n at large frequencies $k > \mu_n$, where we recall that μ_n denotes the frequency cutoff delimiting the frequency atom ϵ^n from below, we shall replace this by the even more natural $U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha, low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha (U_{<0}^{(\phi^n)} \epsilon_0^n)$. Importantly, note that this is now no longer dependent on the choice of a specific frequency $k = 0$. Moreover, we observe that to formulate the right equation we have to pass to the variable $\tilde{\epsilon}_0^n := U_{<0}^{(\phi^n)} \epsilon_0^n$, and more generally $\tilde{\epsilon}_k^n := U_{<k}^{(\phi^n)} \epsilon_k^n$.

Then we shall use the following very natural equation to select our concentration profiles (which in effect is the same for all frequencies and so we shall not indicate a localization here)

$$(7.34) \quad \boxed{\square \tilde{\epsilon} = 2U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha, low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \tilde{\epsilon}.}$$

Observe that the anti-symmetric matrix $U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha, low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger$ is defined purely in terms of the low frequency constituent u^n , over which we already have control.

After these heuristics, we now proceed with the actual selection of the concentration profiles. Here we shall proceed naturally in two steps, first picking the profile at time $t = 0$ for which of course the flow associated with (7.34) is irrelevant, and then picking the temporally unbounded profiles, using the terminology of [16] and [17]. A technical difficulty here consists in identifying a profile which actually maps into the target sphere \mathbb{S}^m . Also, since functions in $\dot{H}_x^1(\mathbb{R}^2)$ are not even distributions, we start by carefully implementing a frequency cutoff away from $-\infty$, and more precisely to the frequency interval $[-K, \infty)$ for some $K \gg 1$. The number K will later on play the role of a parameter which we need to fine tune. Furthermore, we also need to restrict the frequencies from above, below some threshold K_* . This we do by passing to the truncated data $\phi^{n, < K_*}[0]$. Once we obtain bounds for the S norms of the evolution that are uniform in K_* , we shall be able to invoke a simpler perturbative argument just as in the control of the lowest frequency non-atomic part to infer the desired bound.

The following lemma provides a first version of a decomposition into a bounded and an unbounded profile.

Lemma 7.14. *Writing $\phi^{n, < K_*} = u^n + \epsilon^{n, < K_*}$ and passing to a suitable subsequence with respect to n , there exists $p_* \in \mathbb{S}^m$ and for each $K \geq K_0$ and n sufficiently large, a decomposition (all at time $t = 0$)*

$$\begin{aligned} \epsilon^{n, < K_*} + p_* &= \epsilon_{bounded}^{Kn, < K_*} + \epsilon_{unbounded}^{Kn, < K_*} + \eta^{Kn, < K_*}, \\ \partial_t \epsilon^{n, < K_*} &= \gamma_{bounded}^{Kn, < K_*} + \gamma_{unbounded}^{Kn, < K_*} + \zeta^{Kn, < K_*}, \end{aligned}$$

where we have

$$\begin{aligned} \limsup_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} (\|\nabla_x \epsilon_{\text{bounded}}^{Kn, < K_*}\|_{L_x^2}^2 + \|\gamma_{\text{bounded}}^{Kn, < K_*}\|_{L_x^2}^2) &\leq \limsup_{n \rightarrow \infty} \|\nabla_{t,x} \epsilon^{n, < K_*}\|_{L_x^2}, \\ \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \|\nabla_x \eta^{Kn, < K_*}\|_{L_x^2} &= 0, \\ \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \|\zeta^{Kn, < K_*}\|_{L_x^2} &= 0, \end{aligned}$$

as well as the asymptotic orthogonality relations

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{\mathbb{R}^2} (\nabla_x \epsilon_{\text{bounded}}^{Kn, < K_*})^\dagger \cdot \nabla_x \epsilon_{\text{unbounded}}^{Kn, < K_*}(x) dx &= 0, \\ \lim_{n \rightarrow \infty} \int_{\mathbb{R}^2} (\gamma_{\text{bounded}}^{Kn, < K_*})^\dagger \cdot \gamma_{\text{unbounded}}^{Kn, < K_*}(x) dx &= 0. \end{aligned}$$

The components $\epsilon_{\text{bounded}}^{Kn, < K_*}$ map into \mathbb{S}^m and the data pair $(\epsilon_{\text{bounded}}^{Kn, < K_*}, \gamma_{\text{bounded}}^{Kn, < K_*})$ is admissible in the sense that

$$(\epsilon_{\text{bounded}}^{Kn, < K_*})^\dagger \cdot \gamma_{\text{bounded}}^{Kn, < K_*} = 0$$

pointwise. Moreover, the limits $\lim_{K \rightarrow \infty} \nabla_x \epsilon_{\text{bounded}}^{Kn, < K_*} =: \nabla_x \epsilon_{\text{bounded}}^{< K_*}$ and $\lim_{K \rightarrow \infty} \gamma_{\text{bounded}}^{Kn, < K_*} =: \gamma_{\text{bounded}}^{< K_*}$ exist in L_x^2 independently of n with

$$\epsilon_{\text{bounded}}^{< K_*} \in L_x^\infty \cap \dot{H}_x^1, \quad \gamma_{\text{bounded}}^{< K_*} \in L_x^2.$$

We also have the compatibility relation $(\epsilon_{\text{bounded}}^{< K_*})^\dagger \cdot \gamma_{\text{bounded}}^{< K_*} = 0$ pointwise. Furthermore, it holds that $\epsilon_{\text{unbounded}}^{Kn, < K_*} \rightarrow 0$ in L_{loc}^∞ as $n \rightarrow \infty$. Finally, we have that $u^n - p_* \rightarrow 0$ in L_{loc}^∞ as $n \rightarrow \infty$ and for any $R_0 > 0$, we have

$$\begin{aligned} \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \|\chi_{B_{R_0}} \eta^{Kn, < K_*}\|_{L_x^\infty \cap \dot{H}_x^1} &= 0, \\ \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \|\chi_{B_{R_0}} \zeta^{Kn, < K_*}\|_{L_x^2} &= 0. \end{aligned}$$

Proof. The extra parameter K here plays the role of an additional frequency cutoff. Write

$$\epsilon^{n, < K_*} = P_{[-K, \infty)} \epsilon^{n, < K_*} + P_{(-\infty, -K)} \epsilon^{n, < K_*}.$$

The sequence $\{P_{[-K, \infty)} \epsilon^n\}_n$ being bounded in $H_x^1(\mathbb{R}^2)$, passing to a subsequence we may pick a weak limit $\epsilon_{\text{bounded}}^{K, < K_*}$, which is either zero or non-zero. Then upon passing to the subsequence,

$$P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{\text{bounded}}^{K, < K_*}$$

converges weakly toward zero as $n \rightarrow \infty$, and in particular, it converges pointwise toward zero (due to the essentially sharp frequency localization). By letting K run through all positive integers and implementing a Cantor diagonal argument to successive subsequences, we can arrange that $P_{[-K, \infty)} \epsilon_{\text{bounded}}^{\tilde{K}, < K_*} = \epsilon_{\text{bounded}}^{K, < K_*}$ for $\tilde{K} > K$, and thence that $\nabla_x \epsilon_{\text{bounded}}^{K, < K_*}$ converges in L_x^2 as $K \rightarrow \infty$ to, say, $\nabla_x \epsilon_{\text{bounded}}^{< K_*}$ with $\epsilon_{\text{bounded}}^{< K_*} \in \dot{H}_x^1 \cap L_x^\infty$. Similarly, we may assume that $\lim_{K_* \rightarrow \infty} \nabla_x \epsilon_{\text{bounded}}^{< K_*}$ exists in L_x^2 . Now write

$$\phi^{n, < K_*} = u^n + \epsilon_{\text{bounded}}^{K, < K_*} + P_{(-\infty, -K]} \epsilon^{n, < K_*} + (P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{\text{bounded}}^{K, < K_*}).$$

Localizing to a large ball B_R around the origin, notice that (with the error vanishing as $n \rightarrow \infty$)

$$\begin{aligned} u^n|_{B_R} &= p_n + o_{L_x^\infty}(1), \\ \chi_{B_R}(P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{\text{bounded}}^{K, < K_*}) &= o_{L_x^\infty}(1). \end{aligned}$$

It follows that $(\epsilon_{bounded}^{K, < K_*} + P_{(-\infty, -K]} \epsilon^{n, < K_*})|_{B_R} + p_n$ is within a $o_{L_x^\infty}(1)$ neighborhood of \mathbb{S}^m . Passing to a subsequence, we may suppose $p_n \rightarrow p_*$ for some $p_* \in \mathbb{S}^m$, whence $(\epsilon_{bounded}^{K, < K_*} + P_{(-\infty, -K]} \epsilon^{n, < K_*})|_{B_R} + p_*$ is arbitrarily close to \mathbb{S}^m provided n is sufficiently large. Further, observe that

$$\|\epsilon_{bounded}^{K, < K_*}\|_{L_x^\infty(\{R \sim R_*\})} \lesssim (2^K R_*)^{-1} + \|\nabla \epsilon_{bounded}^{K, < K_*}\|_{L_x^2(\{R \sim R_*\})}.$$

Now for a suitable quantity $\delta(K) \rightarrow 0$ as $K \rightarrow \infty$, put $R_* \geq 2^{-K} \delta(K)^{-\frac{1}{2}}$. By convergence of $\nabla \epsilon_{bounded}^{K, < K_*}$ in L_x^2 as $K \rightarrow \infty$, we get from this

$$\|\epsilon_{bounded}^{K, < K_*}\|_{L_x^\infty(\{R > 2^{-K} \delta(K)^{-\frac{1}{2}}\})} \rightarrow 0$$

as $K \rightarrow \infty$. Also, putting

$$\delta(K, K_*) := \limsup_{n \rightarrow \infty} \|\nabla P_{(-\infty, -K]} \epsilon^{n, < K_*}\|_{L_x^2},$$

then indeed we have $\lim_{K \rightarrow \infty} \delta(K, K_*) = 0$, and passing to a subsequence as $K \rightarrow \infty$ we can find $p_1 \in \mathbb{R}^{m+1}$ such that

$$P_{(-\infty, -K]} \epsilon^{n, < K_*}|_{B_{R_*K} \setminus B_{\frac{1}{2}R_*K}} = p_1 + o_{L_x^\infty}(1)$$

as $n, K \rightarrow \infty$, where $R_{*K} := 2^{-K} \delta(K, K_*)^{-\frac{1}{2}}$, since the variation of the function $P_{(-\infty, -K]} \epsilon^{n, < K_*}$ on $B_{R_{*K}} \setminus B_{\frac{1}{2}R_{*K}}$ is $\lesssim \delta(K, K_*)^{\frac{1}{2}}$ as n is very large.

It follows that choosing $\chi_{B_{R_{*K}}}$ as before, the expression

$$\chi_{B_{R_{*K}}} \epsilon_{bounded}^{K, < K_*} + \chi_{B_{R_{*K}}} P_{(-\infty, -K]} \epsilon^{n, < K_*} + (1 - \chi_{B_{R_{*K}}}) \cdot p_1 + p_*$$

converges toward \mathbb{S}^m as $n, K \rightarrow \infty$. Letting Π be the normal projection onto \mathbb{S}^m and labelling

$$\begin{aligned} \tilde{\epsilon}_{bounded}^{Kn, < K_*} &:= \Pi(\chi_{B_{R_{*K}}} \epsilon_{bounded}^{K, < K_*} + \chi_{B_{R_{*K}}} P_{(-\infty, -K]} \epsilon^{n, < K_*} + (1 - \chi_{B_{R_{*K}}}) \cdot p_1 + p_*), \\ \epsilon_{unbounded}^{Kn, < K_*} &:= P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{bounded}^{K, < K_*}, \\ \eta^{nK, < K_*} &:= (1 - \Pi)(\chi_{B_{R_{*K}}} \epsilon_{bounded}^{K, < K_*} + \chi_{B_{R_{*K}}} P_{(-\infty, -K]} \epsilon^{n, < K_*} + (1 - \chi_{B_{R_{*K}}}) \cdot p_1 + p_*) \\ &\quad + (1 - \chi_{B_{R_{*K}}})(P_{(-\infty, -K]} \epsilon^{n1, < K_*} - p_1) + (1 - \chi_{B_{R_{*K}}}) \epsilon_{bounded}^{K, < K_*}, \end{aligned}$$

we infer a representation

$$\epsilon^{n, < K_*} + p_* = \tilde{\epsilon}_{bounded}^{Kn, < K_*} + \epsilon_{unbounded}^{Kn, < K_*} + \eta^{nK, < K_*}$$

with all the properties of the lemma (but with $\tilde{\epsilon}_{bounded}^{Kn, < K_*}$ in place of $\epsilon_{bounded}^{Kn, < K_*}$). We carefully observe that

$$(1 - \chi_{B_{R_{*K}}})(P_{(-\infty, -K]} \epsilon^{n, < K_*} - p_1) + (1 - \chi_{B_{R_{*K}}}) \epsilon_{bounded}^{K, < K_*}$$

is supported outside of B_{R_0} for K large enough, and that

$$(1 - \Pi)(\chi_{B_{R_{*K}}} \epsilon_{bounded}^{K, < K_*} + \chi_{B_{R_{*K}}} P_{(-\infty, -K]} \epsilon^{n, < K_*} + (1 - \chi_{B_{R_{*K}}}) \cdot p_1 + p_*)$$

converges toward zero in the L_x^∞ -norm as $n, K \rightarrow \infty$. It remains to show smallness of the term

$$(1 - \Pi)(\chi_{B_{R_{*K}}} \epsilon_{bounded}^{K, < K_*} + \chi_{B_{R_{*K}}} P_{(-\infty, -K]} \epsilon^{n, < K_*} + (1 - \chi_{B_{R_{*K}}}) \cdot p_1 + p_*).$$

with respect to \dot{H}_x^1 . This is clear by an argument as above, provided we include a cutoff $\chi_{\frac{1}{2}B_{R_{*K}}}^c$ in front. Thus, consider now the term

$$\begin{aligned} &\chi_{\frac{1}{2}B_{R_{*K}}} (1 - \Pi)(\chi_{B_{R_{*K}}} \epsilon_{bounded}^{K, < K_*} + \chi_{B_{R_{*K}}} P_{(-\infty, -K]} \epsilon^{n, < K_*} + (1 - \chi_{B_{R_{*K}}}) \cdot p_1 + p_*) \\ &= \chi_{\frac{1}{2}B_{R_{*K}}} (1 - \Pi)(\epsilon_{bounded}^{K, < K_*} + P_{(-\infty, -K]} \epsilon^{n, < K_*} + p_*) \\ &= \chi_{\frac{1}{2}B_{R_{*K}}} (1 - \Pi)(\phi^{n, < K_*} - (u^n - p_*) - (P_{[-K, \infty)} \epsilon^{n1, < K_*} - \epsilon_{bounded}^{K, < K_*})). \end{aligned}$$

Extending $(1 - \Pi)$ smoothly to all of \mathbb{R}^{m+1} and using the same notation for the global operator, we get on account of

$$(1 - \Pi)(\phi^{n, < K_*}) = 0$$

the following

$$\begin{aligned} & \left\| \nabla_{t,x} \chi_{\frac{1}{2}B_{R_*K}} (1 - \Pi)(\phi^{n, < K_*} - (u^n - p_*) - (P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{bounded}^{K, < K_*})) \right\|_{L_x^2} \\ &= \left\| \nabla_{t,x} \int_0^1 \chi_{\frac{1}{2}B_{R_*K}} \partial_s (1 - \Pi)(\phi^{n, < K_*} - s(u^n - p_*) - s(P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{bounded}^{K, < K_*})) ds \right\|_{L_x^2} \\ &\lesssim \left\| \nabla_{t,x} (\chi_{\frac{1}{2}B_{R_*K}} (u^n - p_*)) \right\|_{L_x^2} + \left\| \nabla_{t,x} (\chi_{\frac{1}{2}B_{R_*K}} (P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{bounded}^{K, < K_*})) \right\|_{L_x^2} \\ &\quad + \left\| (\chi_{\frac{1}{2}B_{R_*K}} (u^n - p_*)) \right\|_{L_x^\infty} + \left\| (\chi_{\frac{1}{2}B_{R_*K}} (P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{bounded}^{K, < K_*})) \right\|_{L_x^\infty}. \end{aligned}$$

All terms at the end are easily seen to converge to 0 as $n \rightarrow \infty$. To see this for the second term, use that

$$\left\| \nabla_x (\chi_{\frac{1}{2}B_{R_*K}} (P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{bounded}^{K, < K_*})) \right\|_{L_x^2} \rightarrow 0$$

as $n \rightarrow \infty$ since $P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{bounded}^{K, < K_*} \rightarrow 0$ weakly as $n \rightarrow \infty$. Also we have

$$\begin{aligned} \left\| \chi_{\frac{1}{2}B_{R_*K}} \nabla_x (P_{[-K, \infty)} \epsilon^{n, < K_*} - \epsilon_{bounded}^{K, < K_*}) \right\|_{L_x^2} &\leq \left\| \chi_{\frac{1}{2}B_{R_*K}} \nabla_x (P_{[-K, K]} \epsilon^{n, < K_*} - P_{(-\infty, K]} \epsilon_{bounded}^{K, < K_*}) \right\|_{L_x^2} \\ &\quad + \left\| \chi_{\frac{1}{2}B_{R_*K}} \nabla_x (P_{(K, \infty)} \epsilon^{n, < K_*} - P_{(K, \infty)} \epsilon_{bounded}^{K, < K_*}) \right\|_{L_x^2}. \end{aligned}$$

Then the first term on the right-hand side converges to 0 as $n \rightarrow \infty$ and the second converges to 0 as $K \rightarrow \infty$ uniformly in n . For the term

$$\left\| \nabla_{t,x} (\chi_{\frac{1}{2}B_{R_*K}} (u^n - p_*)) \right\|_{L_x^2},$$

use the fact that $\lim_{n \rightarrow \infty} \chi_{\frac{1}{2}B_{R_*K}} (u^n - p_*) = 0$ as well as the evacuation to extremely low frequencies for u^n as $n \rightarrow \infty$. We also observe here that since $\nabla_x \epsilon_{bounded}^{Kn, < K_*} \rightarrow \nabla_x \epsilon_{bounded}^{K, < K_*}$ for any large enough n and $K \rightarrow \infty$, we may indeed pick $\epsilon_{bounded}^{K, < K_*}$ to map into \mathbb{S}^m .

As for decomposing the time derivative, let $\gamma_{bounded}^{K, < K_*}$ be a weak limit of $P_{[-K, \infty)} \partial_t \epsilon^{n, < K_*}$ as $n \rightarrow \infty$, and then let

$$\begin{aligned} \gamma_{unbounded}^{Kn, < K_*} &:= P_{[-K, \infty)} \partial_t \epsilon^{n, < K_*} - \gamma_{bounded}^{K, < K_*}, \\ \zeta^{Kn, < K_*} &:= P_{(-\infty, -K]} \partial_t \epsilon^{n, < K_*}. \end{aligned}$$

We have

$$0 = (u^n - p_* + \epsilon_{bounded}^{Kn, < K_*} + \epsilon_{unbounded}^{Kn, < K_*} + \eta^{Kn, < K_*})^\dagger \cdot (\partial_t u^n + \gamma_{bounded}^{K, < K_*} + \gamma_{unbounded}^{Kn, < K_*} + \zeta^{Kn, < K_*}).$$

But then on any bounded set D we have in the pointwise sense that

$$\lim_{n \rightarrow \infty} (u^n - p_* + \epsilon_{unbounded}^{Kn, < K_*})^\dagger \cdot (\partial_t u^n + \gamma_{bounded}^{K, < K_*} + \gamma_{unbounded}^{Kn, < K_*} + \zeta^{Kn, < K_*}) = 0$$

and similarly (on a bounded set D) we have the pointwise limit

$$\lim_{K \rightarrow \infty} (\eta^{Kn, < K_*})^\dagger \cdot (\partial_t u^n + \gamma_{bounded}^{K, < K_*} + \gamma_{unbounded}^{Kn, < K_*} + \zeta^{Kn, < K_*}) = 0.$$

Also, again exploiting the frequency localization, we get

$$\lim_{n \rightarrow \infty} (\epsilon_{bounded}^{Kn, < K_*})^\dagger (\gamma_{unbounded}^{Kn, < K_*} + \zeta^{Kn, < K_*}) = 0.$$

Finally, we conclude that if we put

$$\tilde{\gamma}_{bounded}^{Kn, < K_*} := \gamma_{bounded}^{K, < K_*} - \epsilon_{bounded}^{Kn, < K_*} (\epsilon_{bounded}^{Kn, < K_*})^\dagger \cdot \gamma_{bounded}^{K, < K_*}$$

and then write

$$\partial_t \epsilon^{n, < K_*} = \tilde{\gamma}_{\text{bounded}}^{Kn, < K_*} + \gamma_{\text{unbounded}}^{Kn, < K_*} + \tilde{\zeta}^{Kn, < K_*}$$

with

$$\tilde{\zeta}^{Kn, < K_*} = \zeta^{Kn, < K_*} + \epsilon_{\text{bounded}}^{Kn, < K_*} (\epsilon_{\text{bounded}}^{Kn, < K_*})^\dagger \cdot \gamma_{\text{bounded}}^{Kn, < K_*}.$$

Then we have found the desired representation with $\tilde{\gamma}_{\text{bounded}}^{Kn, < K_*}$ in place of $\gamma_{\text{bounded}}^{Kn, < K_*}$ and $\tilde{\zeta}^{Kn, < K_*}$ in place of $\zeta^{Kn, < K_*}$. \square

Note from the preceding proof that setting now

$$\epsilon_{\text{bounded}}^{Kn, < K_*} = \Pi(\chi_{B_{R_*K}} \epsilon_{\text{bounded}}^{Kn, < K_*} + \chi_{B_{R_*K}} P_{(-\infty, -K]} \epsilon^{n, < K_*} + (1 - \chi_{B_{R_*K}}) \cdot p_1 + p_*),$$

the implied frequency localization from above (up to exponential tails) allows us to conclude the more precise local convergence statement that on any bounded set D we have

$$\lim_{K \rightarrow \infty} \|\epsilon_{\text{bounded}}^{Kn, < K_*} - \epsilon_{\text{bounded}}^{< K_*}\|_{L_x^\infty \cap \dot{H}_x^1(D)} = 0$$

uniformly for all sufficiently large n . Similarly, upon passing to a subsequence if necessary, we have that for any such D

$$\lim_{K \rightarrow \infty} \|\gamma_{\text{bounded}}^{Kn, < K_*} - \gamma_{\text{bounded}}^{< K_*}\|_{L_x^2(D)} = 0$$

for a suitable limit function $\gamma_{\text{bounded}}^{< K_*} \in L_x^2$.

In the preceding we have always kept the *upper frequency bound* K_* fixed. However, we may now increase K_* and only consider subsequences of the sequence of n considered previously. Doing this we may achieve that actually

$$\lim_{K_* \rightarrow \infty} (\nabla_x \epsilon_{\text{bounded}}^{< K_*}, \gamma_{\text{bounded}}^{< K_*})$$

exists in L_x^2 . In what follows we shall use the notation $\epsilon_{\text{bounded}}^{< K_*}[0] := (\epsilon_{\text{bounded}}^{< K_*}, \gamma_{\text{bounded}}^{< K_*})$.

It now remains to reveal the fine structure of the data pair $(\epsilon_{\text{unbounded}}^{Kn, < K_*}, \gamma_{\text{unbounded}}^{Kn, < K_*})$, i.e. to resolve it into profiles. To this end we shall use the key equation (7.34). As we have to work with the gauged variables for each frequency block, we use for each dyadic frequency $k \in \mathbb{Z}$ the data pair

$$(7.35) \quad \tilde{\epsilon}_{\text{unbounded}, k}^{Kn, < K_*}[0] := (U_{< k}^{(\phi^{n, < K_*})} \epsilon_{\text{unbounded}, k}^{Kn, < K_*}, U_{< k}^{(\phi^{n, < K_*})} \gamma_{\text{unbounded}}^{Kn, < K_*} + \partial_t(U_{< k}^{(\phi^{n, < K_*})}) \epsilon_{\text{unbounded}, k}^{Kn, < K_*}).$$

To state the decomposition into concentration profiles for the linear evolution, we first need a precise description of solutions to (7.34).

Lemma 7.15. *Given $\tilde{\epsilon}[0] \in \dot{H}_x^1 \times L_x^2(\mathbb{R}^2; \mathbb{R}^{m+1})$, there exists a unique solution $\tilde{\epsilon}^{(n)}(t, \cdot) \in C^0(\mathbb{R}; \dot{H}_x^1 \times L_x^2(\mathbb{R}^2; \mathbb{R}^{m+1}))$ to (7.34) satisfying uniform bounds*

$$\|\tilde{\epsilon}^{(n)}\|_{S[\mathbb{R}]} \leq C(\|u^n\|_{S[\mathbb{R}]}) \|\tilde{\epsilon}[0]\|_{\dot{H}_x^1 \times L_x^2(\mathbb{R}^2; \mathbb{R}^{m+1})}.$$

The solution $\tilde{\epsilon}^{(n)}(t, \cdot)$ vanishes asymptotically in the sense that given $\gamma > 0$, there exists a decomposition

$$\tilde{\epsilon}^{(n)} = \tilde{\epsilon}_1^{(n)} + \tilde{\epsilon}_2^{(n)}$$

and a time $t_0(\gamma, \|u^n\|_S, \tilde{\epsilon}[0]) > 0$ such that for all n ,

$$\|\tilde{\epsilon}_1^{(n)}\|_{S[\mathbb{R}]} < \gamma, \quad \|\tilde{\epsilon}_2^{(n)}\|_{L_t^\infty L_x^\infty[(-\infty, -t_0] \cup [t_0, \infty)]} < \gamma.$$

Finally, we have asymptotic energy conservation in the sense that

$$\lim_{n \rightarrow \infty} \sup_{t \in \mathbb{R}} \left| \|\nabla_{t,x} \tilde{\epsilon}^{(n)}(t, \cdot)\|_{L_x^2} - \|\tilde{\epsilon}[0]\|_{\dot{H}_x^1 \times L_x^2} \right| = 0.$$

Proof. The a priori bounds follow from the divisibility argument used in the proof of Proposition 6.3. Moreover, we can infer that if $\{c_k\}_{k \in \mathbb{Z}}$ is a sufficiently flat frequency envelope covering the data $\tilde{\epsilon}[0]$, then we obtain

$$\|\tilde{\epsilon}_k^{(n)}\|_{S_k[\mathbb{R}]} \leq D(\|u^n\|_{S[\mathbb{R}]})c_k.$$

The asymptotic decay follows as in the proof of Proposition 9.20 in [17]. Correspondingly, we only turn to the details of the proof of the asymptotic energy conservation. Our reasoning will be closely related to but in certain aspects a much simplified version of the proof of Proposition 9.14 in [17]. Thus, let P_k be slightly modified Littlewood-Paley projections with the property that the corresponding cutoffs $\chi_k(\xi)$ on the frequency side satisfy for all $\xi \neq 0$ that

$$\sum_{k \in \mathbb{Z}} \chi_k(\xi)^2 = 1.$$

In particular, we then have (setting $\tilde{\epsilon}_k = P_k \tilde{\epsilon}$)

$$\sum_{k \in \mathbb{Z}} \|\nabla_{t,x} \tilde{\epsilon}_k(0, \cdot)\|_{L_x^2}^2 = \|\nabla_{t,x} \tilde{\epsilon}(0, \cdot)\|_{L_x^2}^2.$$

In particular, it suffices to prove asymptotic energy conservation for the frequency localized functions $\tilde{\epsilon}_k$. To simplify the notation within this proof, we shall write

$$\mathcal{A}_\alpha \equiv U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha, low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger.$$

Then on localizing (7.34) to dyadic frequency $k \in \mathbb{Z}$, we obtain

$$\square \tilde{\epsilon}_k^{(n)} = 2\mathcal{A}_{\alpha, <k} \partial^\alpha \tilde{\epsilon}_k^{(n)} + F_k^{(n)},$$

where we have $\lim_{n \rightarrow \infty} \sum_k \|F_k^{(n)}\|_{L_t^1 L_x^2}^2 = 0$. We shall now consider the quantity

$$E_k^{(n)}(t) := \frac{1}{2} \int_{\mathbb{R}^2} (|\partial_t \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{0, <k} \tilde{\epsilon}_k^{(n)}|^2 + \sum_{j=1,2} |\partial_j \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{j, <k} \tilde{\epsilon}_k^{(n)}|^2) dx$$

and show that uniformly for all $t \in \mathbb{R}$,

$$(7.36) \quad \lim_{n \rightarrow \infty} E_k^{(n)}(t) = \|\nabla_{t,x} \tilde{\epsilon}_k(0, \cdot)\|_{L_x^2}^2.$$

Then on account of the fact that uniformly for all $t \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} \sum_{k \in \mathbb{Z}} \|\mathcal{A}_{\alpha, <k} \tilde{\epsilon}_k^{(n)}\|_{L_x^2}^2 = 0,$$

the final conclusion of the lemma follows. In order to show (7.36), we differentiate $E_k^{(n)}(t)$ and find

$$\begin{aligned} (E_k^{(n)})'(t) &= \int_{\mathbb{R}^2} (\partial_t \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{0, <k} \partial_t \tilde{\epsilon}_k^{(n)} - \partial_t \mathcal{A}_{0, <k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot (\partial_t \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{0, <k} \tilde{\epsilon}_k^{(n)}) dx \\ &\quad + \sum_{j=1,2} \int_{\mathbb{R}^2} (\partial_j \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{j, <k} \partial_j \tilde{\epsilon}_k^{(n)} - \partial_j \mathcal{A}_{j, <k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot (\partial_j \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{j, <k} \tilde{\epsilon}_k^{(n)}) dx \\ &\equiv I + II. \end{aligned}$$

To simplify things below, we first observe the following schematic vanishing relations

$$\begin{aligned}
\lim_{n \rightarrow \infty} \int_{\mathbb{R}^{1+2}} \partial^\alpha \mathcal{A}_{\alpha, < k} \nabla_{t,x} \tilde{\epsilon}_k^{(n)} \tilde{\epsilon}_k^{(n)} dx dt &= 0, \\
\lim_{n \rightarrow \infty} \int_{\mathbb{R}^{1+2}} (\partial_t \mathcal{A}_{j, < k} - \partial_j \mathcal{A}_{0, < k}) \nabla_{t,x} \tilde{\epsilon}_k^{(n)} \tilde{\epsilon}_k^{(n)} dx dt &= 0, \\
\lim_{n \rightarrow \infty} \int_{\mathbb{R}^{1+2}} (\mathcal{A}_{< k})^2 \nabla_{t,x} \tilde{\epsilon}_k^{(n)} \tilde{\epsilon}_k^{(n)} dx dt &= 0, \\
\lim_{n \rightarrow \infty} \int_{\mathbb{R}^{1+2}} (\nabla_{t,x} \mathcal{A}_{< k}) \mathcal{A}_{< k} (\tilde{\epsilon}_k^{(n)})^2 dx dt &= 0.
\end{aligned}$$

Here one may replace \mathbb{R}^{1+2} by $I \times \mathbb{R}^2$ for any time interval I , the vanishing relations being uniform in I . To see the first of these relations, we write schematically

$$\partial^\alpha \mathcal{A}_{\alpha, < k} = P_{< k} \left(U_{< \mu_n}^{(u^n)} \partial_\alpha u^n \partial^\alpha u^n (U_{< \mu_n}^{(u^n)})^\dagger \right) + \text{cubic terms},$$

where the cubic terms arise upon differentiating $U_{< \mu_n}^{(u^n)}$ or re-expanding $\square u^n$ using the wave maps equation. Then it is straightforward to place the cubic terms into $L_t^1 L_x^M$ using the Strichartz type norms in our S space in conjunction with Bernstein's inequality. In fact, from the definition of u^n , the cubic terms live at frequency $< \mu_n$ up to asymptotically vanishing terms with respect to $L_t^1 L_x^M$. Then one closes by observing that

$$\|P_l(\nabla_{t,x} \tilde{\epsilon}_k^{(n)} \tilde{\epsilon}_k^{(n)})\|_{L_t^\infty L_x^{1+}} \lesssim 2^{-\sigma|l-k|} 2^{-(1-)|k|} \|\tilde{\epsilon}_k^{(n)}\|_{S_k}^2 \lesssim 2^{-\sigma|l-k|} 2^{-(1-)|k|} c_k,$$

where $\{c_k\}_{k \in \mathbb{Z}}$ is a sufficiently flat frequency envelope covering the data $\tilde{\epsilon}[0]$. On the other hand, owing to the null structure, the principal term

$$P_{< k} \left(U_{< \mu_n}^{(u^n)} \partial_\alpha u^n \partial^\alpha u^n (U_{< \mu_n}^{(u^n)})^\dagger \right)$$

can be placed into $L_{t,x}^{\frac{3}{2}+}$, again essentially reduced to frequencies $< \mu_n$, while one uses

$$\|P_l(\nabla_{t,x} \tilde{\epsilon}_k^{(n)} \tilde{\epsilon}_k^{(n)})\|_{L_{t,x}^{3-}} \lesssim 2^{-\sigma|l-k|} 2^{-(0+)|k|} c_k.$$

The second null form

$$\int_{\mathbb{R}^2} (\partial_t \mathcal{A}_{j, < k} - \partial_j \mathcal{A}_{0, < k}) \nabla_{t,x} \tilde{\epsilon}_k^{(n)} \tilde{\epsilon}_k^{(n)} dx$$

is handled similarly, and since we can similarly bound $\|(\mathcal{A}_{< k})^2\|_{L_{t,x}^{\frac{3}{2}+}}$ and $\|(\nabla_{t,x} \mathcal{A}_{< k}) \mathcal{A}_{< k}\|_{L_{t,x}^{\frac{3}{2}+}}$, we also get the remaining vanishing relations.

It now remains to reduce $(E_k^{(n)})'(t)$ to the expressions in the preceding vanishing relations. To this end we first observe that by the anti-symmetry of \mathcal{A}_α it holds that

$$(\mathcal{A}_{0, < k} \partial_t \tilde{\epsilon}_k^{(n)})^\dagger \cdot \partial_t \tilde{\epsilon}_k^{(n)} = 0,$$

and thus for the term I we have

$$I = \int_{\mathbb{R}^2} (\partial_{tt} \tilde{\epsilon}_k^{(n)} - 2\mathcal{A}_{0, < k} \partial_t \tilde{\epsilon}_k^{(n)} - \partial_t \mathcal{A}_{0, < k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot (\partial_t \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{0, < k} \tilde{\epsilon}_k^{(n)}) dx + \text{error},$$

where *error* refers to terms satisfying the above vanishing relations. Next, we integrate by parts in term II to find that

$$\begin{aligned} II &= \sum_{j=1,2} \int_{\mathbb{R}^2} (\partial_t \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{j,<k} \partial_t \tilde{\epsilon}_k^{(n)} - \partial_t \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot (\partial_j \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)}) dx \\ &= - \sum_{j=1,2} \int_{\mathbb{R}^2} (\partial_t \tilde{\epsilon}_k^{(n)})^\dagger \cdot (\partial_{jj} \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{j,<k} \partial_j \tilde{\epsilon}_k^{(n)} - \partial_j \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)}) dx \\ &\quad - \sum_{j=1,2} \int_{\mathbb{R}^2} (\mathcal{A}_{j,<k} \partial_t \tilde{\epsilon}_k^{(n)} + \partial_t \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot (\partial_j \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)}) dx. \end{aligned}$$

Then by the anti-symmetry of \mathcal{A}_α we have the relation $(\mathcal{A}_{j,<k} \partial_t \tilde{\epsilon}_k^{(n)})^\dagger \cdot \partial_j \tilde{\epsilon}_k^{(n)} = -(\partial_t \tilde{\epsilon}_k^{(n)})^\dagger \cdot \mathcal{A}_{j,<k} \partial_j \tilde{\epsilon}_k^{(n)}$, and may write the preceding further as

$$\begin{aligned} II &= - \sum_{j=1,2} \int_{\mathbb{R}^2} (\partial_t \tilde{\epsilon}_k^{(n)})^\dagger \cdot (\partial_{jj} \tilde{\epsilon}_k^{(n)} - 2\mathcal{A}_{j,<k} \partial_j \tilde{\epsilon}_k^{(n)} - \partial_j \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)}) dx \\ &\quad - \sum_{j=1,2} \int_{\mathbb{R}^2} (\partial_t \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot \partial_j \tilde{\epsilon}_k^{(n)} dx + \text{error}. \end{aligned}$$

We further modify the first term on the right-hand side above to obtain

$$\begin{aligned} II &= - \sum_{j=1,2} \int_{\mathbb{R}^2} (\partial_t \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{0,<k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot (\partial_{jj} \tilde{\epsilon}_k^{(n)} - 2\mathcal{A}_{j,<k} \partial_j \tilde{\epsilon}_k^{(n)} - \partial_j \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)}) dx \\ &\quad - \sum_{j=1,2} \int_{\mathbb{R}^2} (\mathcal{A}_{0,<k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot \partial_{jj} \tilde{\epsilon}_k^{(n)} dx - \sum_{j=1,2} \int_{\mathbb{R}^2} (\partial_t \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot \partial_j \tilde{\epsilon}_k^{(n)} dx + \text{error} \\ &= - \sum_{j=1,2} \int_{\mathbb{R}^2} (\partial_t \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{0,<k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot (\partial_{jj} \tilde{\epsilon}_k^{(n)} - 2\mathcal{A}_{j,<k} \partial_j \tilde{\epsilon}_k^{(n)} - \partial_j \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)}) dx \\ &\quad + \sum_{j=1,2} \int_{\mathbb{R}^2} (\partial_j \mathcal{A}_{0,<k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot \partial_j \tilde{\epsilon}_k^{(n)} dx - \sum_{j=1,2} \int_{\mathbb{R}^2} (\partial_t \mathcal{A}_{j,<k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot \partial_j \tilde{\epsilon}_k^{(n)} dx + \text{error}, \end{aligned}$$

where we also used that $(\mathcal{A}_{0,<k} \partial_j \tilde{\epsilon}_k^{(n)})^\dagger \cdot \partial_j \tilde{\epsilon}_k^{(n)} = 0$. Putting things together, we find that

$$\begin{aligned} (E_k^{(n)})'(t) &= \int_{\mathbb{R}^2} (-\square \tilde{\epsilon}_k^{(n)} + 2\mathcal{A}_{\alpha,<k} \partial^\alpha \tilde{\epsilon}_k^{(n)} + \partial^\alpha \mathcal{A}_{\alpha,<k} \tilde{\epsilon}_k^{(n)})^\dagger \cdot (\partial_t \tilde{\epsilon}_k^{(n)} - \mathcal{A}_{0,<k} \tilde{\epsilon}_k^{(n)}) dx \\ &\quad + \sum_{j=1,2} \int_{\mathbb{R}^2} ((\partial_j \mathcal{A}_{0,<k} - \partial_t \mathcal{A}_{j,<k}) \tilde{\epsilon}_k^{(n)})^\dagger \cdot \partial_j \tilde{\epsilon}_k^{(n)} dx + \text{error} \end{aligned}$$

and the asymptotic energy conservation follows by using the equation satisfied by $\tilde{\epsilon}_k^{(n)}$ as well as the above vanishing relations. \square

We now turn to the process of extracting linear concentration profiles with respect to the wave operator

$$\tilde{\square}_{A^n} := \square - 2U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,\text{low}}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha.$$

In the following we use notation borrowed from [1], [17] and [16]. We denote by S_{A^n} the wave propagator associated with $\tilde{\square}_{A^n}$ and shall say that a sequence of data pairs $\tilde{\epsilon}^n[0]: \mathbb{R}^2 \rightarrow \mathbb{R}^{m+1} \times \mathbb{R}^{m+1}$ is *essentially supported at frequency* ~ 1 , provided it holds that

$$\lim_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \|P_{[-R,R]^c} \tilde{\epsilon}^n[0]\|_{\dot{H}_x^1 \times L_x^2} = 0.$$

Given a sequence of data $\{\tilde{\epsilon}^n[0]\}_{n \geq 1}$ essentially supported at frequency ~ 1 and uniformly bounded in $\dot{H}_x^1(\mathbb{R}^2; \mathbb{R}^{m+1}) \times L_x^2(\mathbb{R}^2; \mathbb{R}^{m+1})$, we introduce the set

$$\mathcal{U}_{A^n}(\tilde{\epsilon}^n[0]) := \left\{ V \in L_{t,loc}^2 H_x^1(\mathbb{R}^{1+2}; \mathbb{R}^{m+1}) \cap C_t^1 L_x^2(\mathbb{R}^{1+2}; \mathbb{R}^{m+1}) : \exists \{(t_n, x_n)\}_{n \geq 1} \subset \mathbb{R} \times \mathbb{R}^2 \text{ s.t.} \right. \\ \left. S_{A^n}(\tilde{\epsilon}^n[0])(t + t_n, x + x_n) \rightharpoonup V \right\}.$$

Here the weak limit is in the sense of $L_{t,loc}^2 H_x^1$ and we observe that each such weak limit $V(t, x)$ solves $\square V = 0$ in the sense of distributions. We define

$$\eta_{A^n}(\tilde{\epsilon}^n[0]) := \sup \left\{ E(V) = \int_{\mathbb{R}^2} |\nabla_{t,x} V|^2 dx : V \in \mathcal{U}_{A^n}(\tilde{\epsilon}^n[0]) \right\}.$$

Then the extraction process of the linear concentration profiles is summarized in the following proposition, which we formulate in the context of general data which have the weak frequency localization properties of $\epsilon^n[0]$.

Proposition 7.16. *Let $\{\tilde{\epsilon}^n[0]\}_{n \geq 1}$ be a sequence of radially symmetric data pairs $\mathbb{R}^2 \rightarrow \mathbb{R}^{m+1} \times \mathbb{R}^{m+1}$, which are essentially supported at frequency ~ 1 and satisfy a uniform energy bound*

$$\sup_{n \geq 1} \|\tilde{\epsilon}^n[0]\|_{\dot{H}_x^1 \times L_x^2} \lesssim 1.$$

Upon passing to a suitable subsequence, there exists a sequence of profiles $\tilde{\epsilon}_b[0]$, $b \geq 1$, as well as sequences of time shifts $\{t_b^n\}_{n \geq 1}$, $b \geq 1$, satisfying the divergence relation

$$\lim_{n \rightarrow \infty} |t_b^n - t_{b'}^n| = +\infty, \quad b \neq b',$$

so that we have for any $B \geq 1$,

$$S_{A^n}(\tilde{\epsilon}^n[0])(t, x) = \sum_{b=1}^B S_{A^n(t+t_b^n, \cdot)}(\tilde{\epsilon}_b[0])(t - t_b^n, x) + \tilde{\epsilon}_B^n(t, x)$$

and such that

$$\lim_{B \rightarrow \infty} \limsup_{n \rightarrow \infty} \eta_{A^n}(\tilde{\epsilon}_B^n[0]) = 0.$$

If the data $\tilde{\epsilon}^{K_n, < K_}[0]$ also depend on two parameters K and K_* such that*

$$\lim_{K_{1,2} \rightarrow \infty} \limsup_{K_* \rightarrow \infty} \limsup_{n \rightarrow \infty} \|\tilde{\epsilon}^{K_{1n}, < K_*}[0] - \tilde{\epsilon}^{K_{2n}, < K_*}[0]\|_{\dot{H}_x^1 \times L_x^2} = 0$$

and similarly for K_ , then constructing the corresponding profiles $\tilde{\epsilon}_b^{K, < K_*}$ and the remainder term $\tilde{\epsilon}_B^{K_n, < K_*}$, we may assume that the limits*

$$\lim_{K \rightarrow \infty} \lim_{K_* \rightarrow \infty} \tilde{\epsilon}_b^{K, < K_*}[0]$$

exist in $\dot{H}_x^1 \times L_x^2$, and also that $\lim_{K \rightarrow \infty} \lim_{K_ \rightarrow \infty} \eta_{A^n}(\tilde{\epsilon}_B^{K_n, < K_*}[0])$ exists. Finally, we have asymptotic energy conservation*

$$\lim_{n \rightarrow \infty} \|\nabla_{t,x} (S_{A^n(t+t_b^n, \cdot)}(\tilde{\epsilon}_b[0]))(0 - t_b^n, \cdot)\|_{L_x^2} = \|\tilde{\epsilon}_b[0]\|_{\dot{H}_x^1 \times L_x^2}, \quad b \geq 1,$$

as well as the asymptotic orthogonality relation

$$\|\tilde{\epsilon}^n[0]\|_{\dot{H}_x^1 \times L_x^2}^2 = \sum_{b=1}^B \|\tilde{\epsilon}_b[0]\|_{\dot{H}_x^1 \times L_x^2}^2 + \|\tilde{\epsilon}_B^n[0]\|_{\dot{H}_x^1 \times L_x^2}^2 + o(1) \quad \text{as } n \rightarrow \infty.$$

Proof. Using Lemma 7.15, the proof proceeds in direct analogy to the proof of Proposition 7.11 in [17] or to the proof of Lemma 9.23 in [16]. \square

For each dyadic frequency $k \in \mathbb{Z}$ we now consider the sequence $\{\tilde{\epsilon}_{unbounded,k}^{Kn,<K_*}[0]\}_{n \geq 1}$ defined further above in (7.35) and extract concentration profiles $\tilde{\epsilon}_{b,k}^{K,<K_*}[0]$, $b \geq 1$. Applying Proposition 7.16 each time and passing to a suitable subsequence, we may assume that we obtain the same time scales $\{t_b^n\}_{n \geq 1}$. Importantly, by the construction of $\{\tilde{\epsilon}_{unbounded,k}^{Kn,<K_*}[0]\}_{n \geq 1}$ in Lemma 7.14, the extracted concentration profiles $\tilde{\epsilon}_{b,k}^{K,<K_*}[0]$, $b \geq 1$, must all be temporally unbounded, i.e. we have $\lim_{n \rightarrow \infty} |t_b^n| = +\infty$ for every $b \geq 1$. In order to simplify the notation we set

$$\tilde{\epsilon}_{b,k}^{Kn,<K_*}(t, \cdot) := S_{A^n(t+t_b^n, \cdot)}(\tilde{\epsilon}_{b,k}^{K,<K_*}[0])(t - t_b^n, \cdot).$$

Moreover, for later reference, we observe that due to our definition (7.35) and simple frequency considerations, we have that $\tilde{\epsilon}_{b,k}^{Kn,<K_*}$ is essentially supported at frequency $\sim 2^k$ in the sense that

$$\sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_{k \in \mathbb{Z}} \|P_{k+a} \tilde{\epsilon}_{b,k}^{Kn,<K_*}\|_{S_{k+a}}^2 \right)^{\frac{1}{2}} \lesssim \|\tilde{\epsilon}_b^{K,<K_*}[0]\|_{\dot{H}_x^1 \times L_x^2}$$

for an absolute constant $\sigma > 0$.

Then it is natural to make the following ansatz for each dyadic frequency $k \in \mathbb{Z}$,

$$(7.37) \quad \phi_k^{n,<K_*} = u_k^n + \epsilon_{bounded,k}^{<K_*} + (U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) + \eta_{(k)}^{Kn,<K_*}$$

and without the frequency localization

$$(7.38) \quad \phi^{n,<K_*} = u^n - p_* + \epsilon_{bounded}^{<K_*} + \sum_{k \in \mathbb{Z}} (U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) + \eta^{Kn,<K_*}.$$

Here, $\epsilon_{bounded}^{<K_*}$ of course refers to the wave maps evolution of the data $\epsilon_{bounded}^{<K_*}[0] = (\epsilon_{bounded}^{<K_*}, \gamma_{bounded}^{<K_*})$ constructed in the preceding Lemma 7.14.

There is a small subtlety in (7.37) on account of the fact that the expression

$$(U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right)$$

is not necessarily localized to frequency $\sim 2^k$ but only up to exponentially decaying tails, and hence the same applies to $\eta_{(k)}^{Kn,<K_*}$, whence the different notation. Note, however, that by definition we have

$$\begin{aligned} (U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) + \eta_{(k)}^{Kn,<K_*} &= \tilde{P}_k \left((U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) + \eta_{(k)}^{Kn,<K_*} \right) \\ &= \tilde{P}_k \left((U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right) + \tilde{P}_k \eta_{(k)}^{Kn,<K_*} \end{aligned}$$

for a suitable localizer \tilde{P}_k , and this is what we shall substitute in frequency localized terms. Then we can also replace (7.38) by

$$(7.39) \quad \phi^{n,<K_*} = u^n - p_* + \epsilon_{bounded}^{<K_*} + \sum_{k \in \mathbb{Z}} \tilde{P}_k \left((U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right) + \sum_{k \in \mathbb{Z}} \tilde{P}_k \eta_{(k)}^{Kn,<K_*}.$$

More precisely, this ansatz is natural to make as long as we are in the regime where the $\epsilon_{bounded}^{<K_*}$ is truly nonlinear, while in its asymptotic regime we would want to replace it by a solution to (7.34). Unfortunately, by comparison to [17] and [16], it appears harder here to make a good *global* ansatz for the solution, as the previous Lemma 7.14 has already indicated how delicate the correct choice of the bounded profile at time $t = 0$ was. In fact, we expect the correct choice of bounded profile at the next concentration time $t = t_1^n$

(assuming, as we may, that $t_1^n \ll t_2^n \ll \dots$) to delicately hinge on $\phi^{n, < K_*}$ at that time. Our way out of this shall be a careful inductive procedure, first controlling the solution (in terms of the $\eta^{Kn, < K_*}$) on the time slice $[0, t_1^n - T_{1*}]$ for some sufficiently large but finite T_{1*} (and as usual picking n large enough), and then delicately modifying the ansatz (7.39) to track the solution on $[t_1^n - T_{1*}, t_2^n - T_{2*}]$ and so on. Here a crucial point shall be that *our choice of B shall be rather simple*, and in fact only hinge on $\|u^n\|_{S[\mathbb{R}]}$. This should be compared with the cruder arguments in [17] and [16], where the choice of B hinges on fine properties such as S norm bounds of the profiles and their scattering behavior. In our situation, in light of the poor perturbation theory and the fact that we do not even know the later nonlinear profiles, we could not possibly define B in this manner.

We call the expressions

$$\epsilon_{bounded}^{< K_*}[0], \quad \sum_{k \in \mathbb{Z}} \tilde{P}_k((U_{< k}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{b,k}^{Kn, < K_*})[0], \quad b \geq 1,$$

the *profiles* (all evaluated at time $t = 0$). We note that the energy of these expressions as $n \rightarrow \infty$ is well-defined. Also, it follows from Lemma 7.14 that we have the asymptotic orthogonality relation

$$\begin{aligned} & \|\nabla_{t,x} \phi^{n, < K_*}(0, \cdot)\|_{L_x^2}^2 \\ (7.40) \quad &= \|\nabla_{t,x} u^n(0, \cdot)\|_{L_x^2}^2 + \|\nabla_{t,x} \epsilon_{bounded}^{< K_*}(0, \cdot)\|_{L_x^2}^2 + \sum_{b=1}^B \|\nabla_{t,x} \sum_{k \in \mathbb{Z}} \tilde{P}_k((U_{< k}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{b,k}^{Kn, < K_*})(0, \cdot)\|_{L_x^2}^2 \\ &+ \|\nabla_{t,x} \sum_{k \in \mathbb{Z}} \tilde{P}_k((U_{< k}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{B,k}^{Kn, < K_*})(0, \cdot)\|_{L_x^2}^2 + o(1), \end{aligned}$$

where the error vanishes asymptotically as $K, n \rightarrow \infty$. The following theorem is the key result of this subsection.

Theorem 7.17. *Assume that the profiles all have asymptotically (as $K, K_* \rightarrow \infty$) energy strictly less than E_{crit} . In particular, this is the case when there are at least two profiles present for K, K_* large enough. Then the data $\phi^{n, < K_*}[0]$ can be evolved globally in time and the resulting solution $\phi^{n, < K_*}$ satisfies uniformly for large K_* and n the bound*

$$(7.41) \quad \|\phi^{n, < K_*}\|_S < \infty.$$

In fact, given $\delta_2 > 0$, there exists $B = B(\|u^n\|_S, \delta_2)$ such that there are profiles $\tilde{\epsilon}_{bounded}^{n, < K_*}$ and $\tilde{\epsilon}_b^{Kn, < K_*}$, $b \geq 1$, satisfying for K, K_*, n sufficiently large

$$E[\tilde{\epsilon}_{bounded}^{n, < K_*}] < E[\epsilon_{bounded}^{< K_*}] + \delta_2, \quad E[\tilde{\epsilon}_b^{Kn, < K_*}] < E[\tilde{\epsilon}_b^{Kn, < K_*}] + \delta_2, \quad b \geq 1,$$

and also

$$\|\tilde{\epsilon}_{bounded}^{n, < K_*}\|_S < \infty, \quad \|\tilde{\epsilon}_b^{Kn, < K_*}\|_S < \infty, \quad b \geq 1,$$

uniformly for n, K_* large and such that

$$\phi^{n, < K_*} = u^n + \tilde{\epsilon}_{bounded}^{n, < K_*} + \sum_{b=1}^B \tilde{\epsilon}_b^{Kn, < K_*} + \sum_{k \in \mathbb{Z}} (U_{< k}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{B,k}^{Kn, < K_*} + \tilde{\eta}$$

with $\|\tilde{\eta}\|_S < \delta_2$. The profiles $\tilde{\epsilon}_{bounded}^{n, < K_*}$ and $\tilde{\epsilon}_b^{Kn, < K_*}$ coincide with the nonlinear profile $\epsilon_{bounded}^{< K_*}$, respectively certain nonlinear profiles $\tilde{\epsilon}_b^{< K_*}$ near $t = 0$, respectively near $t = t_b^n$, and are of the form

$$\tilde{\epsilon}_b^{Kn, < K_*} = \sum_{k \in \mathbb{Z}} (U_{< k}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{b,k}^{Kn, < K_*},$$

where the $\tilde{\epsilon}_{b,k}^{Kn, < K_*}$ are solutions to (7.34) far away from these times. Also, $\tilde{\epsilon}_{B,k}^{Kn, < K_*}$ is a solution to (7.34).

Remark 7.18. We emphasize that the bound (7.41) on the S norms of the evolutions $\phi^{n,<K_*}$, which holds uniformly for all sufficiently large K_* and n , implies via the high-frequency perturbation Lemma 7.4 that the data $\Pi_{\leq \mu_n^1 + \log(R_n)} \phi^n[0]$ can be globally evolved with uniform S norm bounds for all sufficiently large n . This conclusion is key for the next step in our induction on frequency process in the next subsection.

Before we begin with the proof of Theorem 7.17 we first address some technical issues.

7.4.2. *Technical remarks about multilinear estimates using the decomposition (7.39).* In the sequel, we shall estimate the terms on the right hand side of the basic gauged wave equation

$$(7.42) \quad \square(U_{<k}^{(\phi^{n,<K_*})} \overline{\phi_k^{n,<K_*}}) = U_{<k}^{(\phi^{n,<K_*})} F_k(\phi^{n,<K_*}, \nabla_{t,x} \phi^{n,<K_*})$$

or minor variations thereof by inserting the decompositions (7.39) and exploiting a priori bounds on u^n , $\epsilon_{<K_*}^{<K_*}$, $\tilde{\epsilon}_{b,k}^{Kn,<K_*}$ and $\tilde{\epsilon}_{B,k}^{Kn,<K_*}$, where for the last two expressions we of course need control over all frequencies $k \in \mathbb{Z}$.

We recall that the structure of the renormalized nonlinearity $F_k(\cdot, \cdot)$ was carefully analyzed in Proposition 5.1. For what follows it will be useful to introduce the notion of “perturbative factors” and “non-perturbative factors” in the multilinear expressions constituting the nonlinearity $F_k(\cdot, \cdot)$. We call an input of a multilinear expression in $F_k(\cdot, \cdot)$ a “non-perturbative factor” if it can only be estimated in $L_t^\infty L_x^\infty$ in order to place the whole multilinear expression into $L_t^1 L_x^2$. These “non-perturbative factors” can only occur in the quintilinear expressions¹ $Q_k(\phi)$ in (5.32) as well as in (5.34) and in (5.35). All other inputs of any multilinear term in $F_k(\cdot, \cdot)$ will be referred to as “perturbative factors”.

Then we will be facing the technical difficulty that the “gauged terms”

$$(U_{<k}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{b,k}^{Kn,<K_*}, \quad (U_{<k}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{B,k}^{Kn,<K_*}$$

again involve the a priori uncontrolled function $\phi^{n,<K_*}$, although we of course have a priori control over $\|(U_{<k}^{(\phi^{n,<K_*})})^\dagger\|_{L_{t,x}^\infty}$. Also, observe that the two preceding expressions actually are not localized sharply to frequency $\sim 2^{\tilde{k}}$, but only up to exponentially decaying tails. To deal with these issues we formulate

Lemma 7.19. Let $\{c_k^{(b)}\}_{k \in \mathbb{Z}}$ be a sufficiently flat frequency envelope covering $\sum_{\tilde{k}} 2^{\sigma|k-\tilde{k}|} \|P_k(\tilde{\epsilon}_{b,\tilde{k}}^{Kn,<K_*})\|_{S_k}$ for some small $\sigma > 0$. Then substituting

$$\sum_{\tilde{k}} P_{\tilde{k}}((U_{<\tilde{k}}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{b,\tilde{k}}^{Kn,<K_*})$$

for $\phi^{n,<K_*}$ for the “perturbative factors” in the terms $F_k(\phi^{n,<K_*}, \nabla_{t,x} \phi^{n,<K_*})$, and calling the resulting functions G_k , we get under the bootstrap assumption $\|\sum_k \tilde{P}_k \eta_{(k)}^{Kn,<K_*}\|_S \lesssim 1$ the bound

$$\|G_k\|_{L_t^1 L_x^2} \leq C(\|u^n\|_S, \|\epsilon_{<K_*}^{<K_*}\|_S, \sum_k (\sum_{\tilde{b}} c_k^{(\tilde{b})})^2, \sum_k (\sum_{\tilde{k}} 2^{\sigma|k-\tilde{k}|} \|P_k \tilde{\epsilon}_{B,\tilde{k}}^{Kn,<K_*}\|_S)^2) c_k^{(b)}.$$

Thus, one can estimate G_k as if the factors were just $P_k \sum_{\tilde{k}} \tilde{\epsilon}_{b,\tilde{k}}^{Kn,<K_*}$. Similar inequalities can be obtained when one or more of the “perturbative factors” in $F_k(\cdot, \cdot)$ are occupied by $\sum_{\tilde{k}} P_{\tilde{k}}((U_{<\tilde{k}}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{b,\tilde{k}}^{Kn,<K_*})$ while other “perturbative factors” are occupied by one or more of the remaining terms in (7.39).

Proof. We illustrate it by considering the specific term

$$F_k(\phi, \nabla_{t,x} \phi) = \sum_{k_1 < k} \sum_{k_2 < k_1 - 10} \phi_{k_1} \partial_\alpha \phi_{k_2}^\dagger \partial^\alpha \phi_k.$$

¹An example of a “non-perturbative factor” is the input $\phi_{\leq k-10}^{(1)}$ in the following quintilinear expression $P_k(\phi_{\leq k-10}^{(1)} \sum_{k_1=k_2+O(1)} P_{k_1}(\phi_{\leq k_1-10}^{(2)} \sum_{k_3=k_4+O(1)} (\partial_\alpha \phi_{k_3}^{(3)})^\dagger \partial^\alpha \phi_{k_4}^{(4)})^\dagger \phi_{k_2}^{(5)})$ arising in the term $Q_k(\phi)$ in (5.32).

Thus we have to bound the expression

$$(7.43) \quad \sum_{k_1 < k} \sum_{k_2 < k_1 - 10} P_{k_1}((U_{< k_1}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{b, k_1}^{Kn, < K_*}) \partial_\alpha P_{k_2}((U_{< k_2}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{b, k_2}^{Kn, < K_*})^\dagger \partial^\alpha P_k((U_{< k}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{b, k}^{Kn, < K_*}).$$

As usual, to simplify things, we reduce as we may to the case $k = 0$. To begin with, we infer the general bound

$$(7.44) \quad 2^{\frac{k}{3}} \|P_k((U_{< k}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{b, k}^{Kn, < K_*})\|_{L_t^3 L_x^\infty} \lesssim c_k^{(b)},$$

where the implied constant on the right is like the constant $C(\cdot)$ in the statement of the lemma. To see this, write

$$(7.45) \quad \begin{aligned} P_k((U_{< k}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{b, k}^{Kn, < K_*}) &= P_k(P_{< k-10}(U_{< k}^{(\phi^{n, < K_*})})^\dagger P_{[k-10, k+10]} \tilde{\epsilon}_{b, k}^{Kn, < K_*}) \\ &\quad + P_k(P_{[k-10, k+10]}(U_{< k}^{(\phi^{n, < K_*})})^\dagger P_{< k-10} \tilde{\epsilon}_{b, k}^{Kn, < K_*}) \\ &\quad + P_k(P_{\geq k-10}(U_{< k}^{(\phi^{n, < K_*})})^\dagger P_{\geq k-10} \tilde{\epsilon}_{b, k}^{Kn, < K_*}). \end{aligned}$$

The desired bound for the first term on the right is immediate. The bound for the third term on the right is also immediate, using

$$\|P_{\geq k-10}(U_{< k}^{(\phi^{n, < K_*})})^\dagger\|_{L_t^\infty L_x^2} \lesssim 2^{-k} \|\nabla_{t,x} \phi^{n, < K_*}\|_{L_t^\infty L_x^2}$$

and Bernstein's inequality

$$\begin{aligned} \|P_k(P_{\geq k-10}(U_{< k}^{(\phi^{n, < K_*})})^\dagger P_{\geq k-10} \tilde{\epsilon}_{b, k}^{Kn, < K_*})\|_{L_t^3 L_x^\infty} &\lesssim \sum_{k_1 > k-10} 2^k \|P_{k_1}(U_{< k}^{(\phi^{n, < K_*})})^\dagger\|_{L_t^\infty L_x^2} \|P_{k_1} \tilde{\epsilon}_{b, k}^{Kn, < K_*}\|_{L_t^3 L_x^\infty} \\ &\lesssim 2^k \sum_{k_1 > k-10} 2^{-\frac{4}{3}k_1} 2^{-\sigma|k-k_1|} c_k^{(b)} \lesssim 2^{-\frac{k}{3}} c_k^{(b)}. \end{aligned}$$

For the second term on the right, we expand $P_{[k-10, k+10]}(U_{< k}^{(\phi^{n, < K_*})})^\dagger$ schematically into

$$-2^{-k} P_k((U_{< k}^{(\phi^{n, < K_*})})^\dagger \sum_{k_1 < k} \nabla_x (\phi_{< k_1} \phi_{k_1}^\dagger - \phi_{k_1} \phi_{< k_1}^\dagger) (U_{< k}^{(\phi^{n, < K_*})})^\dagger).$$

Then, depending on where the derivative ∇_x lands, re-expand either $\nabla_x \phi_{< k_1}$ or $\nabla_x \phi_{k_1}$ using (7.39), and repeat application of the above trichotomy. Then using that

$$\|P_{< k-10} \tilde{\epsilon}_{b, k}^{Kn, < K_*}\|_{L_t^M L_x^\infty} \lesssim 2^{-\frac{k}{M}} c_k^{(b)}$$

for M large enough, one obtains the desired bound (7.44) after finitely many steps, using Hölder's inequality.

Now back to (7.43), consider first the case where all derivatives fall on a $\tilde{\epsilon}$ -factor, which is the expression

$$(7.46) \quad \sum_{k_1 < 0} \sum_{k_2 < k_1 - 10} P_{k_1}((U_{< k_1}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{b, k_1}^{Kn, < K_*}) P_{k_2}((U_{< k_2}^{(\phi^{n, < K_*})})^\dagger \partial_\alpha \tilde{\epsilon}_{b, k_2}^{Kn, < K_*})^\dagger P_0((U_{< 0}^{(\phi^{n, < K_*})})^\dagger \partial^\alpha \tilde{\epsilon}_{b, 0}^{Kn, < K_*}).$$

Then if in each of the expressions $P_{k_i}(\dots)$, $i = 1, 2$ and $P_0(\dots)$, we have a low-high interaction, we can simply bound the factors $(U_{< \cdot}^{(\phi^{n, < K_*})})^\dagger$ in $L_{t,x}^\infty$ and the estimate follows. On the other hand, if at least one of the first two of these expressions has a high-low or high-high interaction, we place both factors inside it into $L_t^3 L_x^\infty$, using the observation before, and then also place the remaining low frequency term $P_{k_j}(\dots)$ into $L_t^3 L_x^\infty$, while we simply place $P_0(\dots)$ into $L_t^\infty L_x^2$. In case of a high-low or high-high interaction inside $P_0(\dots)$, we place the lower frequency term into $L_t^3 L_x^\infty$ and the remaining low frequency terms $P_{k_j}(\dots)$, $j = 1, 2$, into $L_t^3 L_x^\infty$, while the high frequency term inside $P_0(\dots)$ gets placed into $L_t^\infty L_x^2$.

If at least one derivative ∂_α falls on a term $(U_{< \cdot}^{(\phi^{n, < K_*})})^\dagger$, say inside $P_{k_2}(\dots)$, one can again close by placing three low frequency terms into $L_t^3 L_x^\infty$ as long as there is a low-high or high-high interaction inside $P_{k_2}(\dots)$.

In case of a high-low interaction, one uses re-expansion of $\partial_\alpha(U_{<k_2}^{(\phi^{n,<K_*})})^\dagger$, using the equation for $U_{<k_2}^{(\phi^{n,<K_*})}$ and (7.39), similar to the proof of (7.44). \square

7.4.3. An important bilinear estimate. In the proof of Theorem 7.17, an important role shall be played by the fact that certain source terms which are multilinear expressions whose factors are all essentially supported at frequency ~ 1 , but which enjoy smallness of their L_x^∞ -norm, have small square-summed (over frequencies) $L_t^1 L_x^2$ -norm. This smallness can be gained on account of our delicate bilinear null form estimate (4.1) which allows us to estimate the null form $\partial_\alpha \phi \partial^\alpha \psi$ “below” the $L_{t,x}^2$ -space, in fact in $L_{t,x}^{\frac{3}{2}}$. This crucial fact was first pointed out by Klainerman-Machedon [8] provided both factors are free waves. We note that the corresponding improvements without the radially assumption are due to Bourgain [2], Wolff [41] and Tao [35] and played an important role in [17]. We quickly illustrate how to use this result with the following

Lemma 7.20. *Let $\{\epsilon^{(n)}\}_{n \geq 1}$ be a family of functions essentially localized to frequency ~ 1 and with uniform bounds*

$$\sup_{n \geq 1} \|\epsilon^{(n)}\|_S \lesssim 1$$

as well as the pointwise decay $\lim_{n \rightarrow \infty} \|P_k \nabla_{t,x} \epsilon^{(n)}\|_{L_{t,x}^\infty} = 0$ for all $k \in \mathbb{Z}$. Then denoting by $F_k(\epsilon^{(n)}, \nabla_{t,x} \epsilon^{(n)})$ any one of the source terms occurring in our equation

$$\square(U_{<k}^{(\phi)} \bar{\phi}_k) = U_{<k}^{(\phi)} F_k(\phi, \nabla_{t,x} \phi),$$

we have

$$\lim_{n \rightarrow \infty} \sum_{k \in \mathbb{Z}} \|F_k(\epsilon^{(n)}, \nabla_{t,x} \epsilon^{(n)})\|_{L_t^1 L_x^2}^2 = 0.$$

Proof. First, we may reduce the output frequency k as well as the frequencies of all perturbative factors in $F_k(\cdot, \cdot)$ to size $O(1)$ due to the essential frequency localization of $\epsilon^{(n)}$. Then we consider a generic expression

$$\epsilon_{k_1}^{(n)} \partial_\alpha \epsilon_{k_2}^{(n)} \partial^\alpha \epsilon_{k_3}^{(n)}, \quad k_j = O(1).$$

We have

$$\|\epsilon_{k_1}^{(n)} \partial_\alpha \epsilon_{k_2}^{(n)} \partial^\alpha \epsilon_{k_3}^{(n)}\|_{L_t^1 L_x^2} \lesssim \|\epsilon_{k_1}^{(n)}\|_{L_t^{2+} L_x^\infty} \|\partial_\alpha \epsilon_{k_2}^{(n)} \partial^\alpha \epsilon_{k_3}^{(n)}\|_{L_{t,x}^{2-}},$$

where the implied constant depends on $k_{1,2,3}$ and we used Bernstein’s inequality to pass from $L_t^1 L_x^2$ to $L_t^1 L_x^{2-}$. Further, by interpolation and the null form estimate (4.1) we have

$$\|\partial_\alpha \epsilon_{k_2}^{(n)} \partial^\alpha \epsilon_{k_3}^{(n)}\|_{L_{t,x}^{2-}} \lesssim \|\partial_\alpha \epsilon_{k_2}^{(n)} \partial^\alpha \epsilon_{k_3}^{(n)}\|_{L_{t,x}^{\frac{3}{2}}}^{\frac{2}{3}} \|\partial_\alpha \epsilon_{k_2}^{(n)} \partial^\alpha \epsilon_{k_3}^{(n)}\|_{L_{t,x}^{\frac{4}{3}}}^{\frac{1}{3}} \lesssim \prod_{j=2,3} \|\epsilon_{k_j}^{(n)}\|_{S_{k_j}^{\frac{3}{2}}}^{\frac{2}{3}} \|\partial_\alpha \epsilon_{k_2}^{(n)} \partial^\alpha \epsilon_{k_3}^{(n)}\|_{L_{t,x}^\infty}^{\frac{1}{3}}.$$

Then the last factor at the end vanishes asymptotically due to our assumption, which yields the claim. \square

The preceding lemma, or simple variations thereof, will play an important role in Step 2 of Stage 1 of the proof of Theorem 7.17 below, and more specifically, in the proof of Lemma 7.22 on an “essentially unbounded” time interval. Working on a bounded time interval as in Step 1 below, one can use much cruder bounds, as for example in (ii) of Step 1 below.

7.4.4. The proof of Theorem 7.17. We now use the preceding preparations to provide

Proof of Theorem 7.17. The strategy shall be to understand the bound as well as the profile decomposition on the interval $[0, t_1^n - T_*]$ and then reveal in a second step how to adjust the profile decomposition to cover also the interval $[t_1^n - T_*, t_2^n - T_{1*}]$ for sufficiently large T_* and T_{1*} as well as for sufficiently large parameters K and n . This will reveal that ensuring $\eta_{A^n}(\epsilon_B^{K^n, < K_*}[0]) < \delta_*(\|u^n\|_S, \delta_2)$ is indeed enough to hold the error term $\tilde{\eta}^{K^n, < K_*} < \delta_2$ even as the process gets continued beyond time t_2^n . At every step, the values of K and n may have to be increased, but since there is only an a priori bounded number of steps, this process will not

diverge. Finally, by their construction, the profiles $\tilde{\epsilon}_{bounded}^{n,<K_*}$ and $\tilde{\epsilon}_b^{Kn,<K_*}$ will satisfy uniform S norm bounds with respect to n , K and K_* sufficiently large. We now proceed in two stages:

Stage 1: *Control of the solution on the first time slice $[0, t_1^n - T_*]$ for sufficiently large T_* .* In turn, we distinguish between the region where $\epsilon_{bounded}^{<K_*}$ is nonlinear and the region where it scatters:

Step 1: *Here we control the solution on any sufficiently large but fixed finite time interval starting at $t = 0$:*

Lemma 7.21. *Given any $R > 0$ (which we will eventually choose depending on the properties of $\epsilon_{bounded}^{<K_*}$) and arbitrary $\delta_3 > 0$, there are K and n sufficiently large such that defining $\sum_k \tilde{P}_k \eta_{(k)}^{Kn,<K_*}$ via (7.39) for a fixed choice of B , we have uniformly for all large K_* that*

$$\left\| \sum_k \tilde{P}_k \eta_{(k)}^{Kn,<K_*} \right\|_{S([0,R])} < \delta_3.$$

In particular, we can achieve that $\left\| \sum_k \tilde{P}_k \eta_{(k)}^{Kn,<K_} \right\|_{S([0,R])} \ll \delta_2$, where $\delta_2 > 0$ is the constant prescribed in the statement of Theorem 7.17.*

Observe that this holds independently of the B chosen in (7.38). The choice of B will be important for the control later on and will be done depending only on the size of $\|u^n\|_S$ and δ_2 . For simplicity of notation, we shall from now on write $\eta^{Kn,<K_*} \equiv \sum_k \tilde{P}_k \eta_{(k)}^{Kn,<K_*}$ and $\eta_k^{Kn,<K_*} \equiv P_k \sum_{\tilde{k}} \tilde{P}_{\tilde{k}} \eta_{(\tilde{k})}^{Kn,<K_*}$.

Proof of Lemma 7.21. We recall the frequency localized wave equation for $\phi^{n,<K_*}$, written schematically as

$$\square \left(U_{<k}^{(\phi^{n,<K_*})} \overline{\phi_k^{n,<K_*}} \right) = U_{<k}^{(\phi^{n,<K_*})} F_k(\phi^{n,<K_*}, \nabla_{t,x} \phi^{n,<K_*}).$$

For technical reasons, specifically the ability to pass from $\overline{\phi_k^{n,<K_*}}$ to $\phi_k^{n,<K_*}$, we shall in fact replace the former quantity by

$$\begin{aligned} \overline{\phi_k^{n,<K_*,R,M}} &:= \phi_k^{n,<K_*} + P_k \left(\phi_{<k-10}^{n,<K_*} \sum_{k_{1,2} > k + \log R} (\phi_{k_1}^{n,<K_*})^\dagger \phi_{k_2}^{n,<K_*} \right) \\ &+ P_k \left(\phi_{<k-10 \log R}^{n,<K_*} \sum_{k_{1,2} \in [k + \log M, k + \log R]} (\phi_{k_1}^{n,<K_*})^\dagger \phi_{k_2}^{n,<K_*} \right) \\ &+ P_k \left(\phi_{<k-10 \log R}^{n,<K_*} \sum_{k < k_{1,2} \leq k + \log M} \chi_{\{r \geq 2^{-k} M\}} (\phi_{k_1}^{n,<K_*})^\dagger \phi_{k_2}^{n,<K_*} \right) \\ &+ P_k \left(\sum_{k_1 < k - \log R} \phi_{k_1-10}^{n,<K_*} \nabla (\phi_{k_1}^{n,<K_*})^\dagger \phi_k^{n,<K_*} \right) \\ &+ P_k \left(\sum_{k_1 \in [k - \log R, k - \log M]} \phi_{k_1-10 \log R}^{n,<K_*} \nabla (\phi_{k_1}^{n,<K_*})^\dagger \phi_k^{n,<K_*} \right) \\ &+ P_k \left(\sum_{k_1 \in [k - \log M, k]} \chi_{\{r \geq 2^{-k} M\}} \phi_{k_1-10 \log R}^{n,<K_*} \nabla (\phi_{k_1}^{n,<K_*})^\dagger \phi_k^{n,<K_*} \right). \end{aligned}$$

Here the parameter M will be chosen large depending on the size of $\|\epsilon_{bounded}^{<K_*}\|_S$ and $\|u^n\|_S$, while the parameter R shall be chosen large depending on more subtle properties of $\epsilon_{bounded}^{<K_*}$ such as its energy distribution among frequencies. In particular, we may assume $R \gg M$. We also need to modify the matrices $U_{<k}^{(\phi^{n,<K_*})}$ in order to render them “less dependent” on $\eta^{Kn,<K_*}$. Specifically, we replace them by $U_{<k}^{(\phi^{n,<K_*},R)}$, which is

constructed analogously to $U_{<k}^{(\phi^n, <K_*)}$ but with the matrix B_h in its construction (5.27) modified to

$$B_h^{(R)} := \frac{d}{dh} \sum_{\ell \in \mathbb{Z}} \eta(h - 10 \log R - \ell) (\phi_{\leq \ell-10} \phi_\ell^\dagger - \phi_\ell \phi_{\leq \ell-10}^\dagger) \\ + \frac{d}{dh} \sum_{\ell \in \mathbb{Z}} \chi_{\{r \geq 2^{-\ell} R^{40}\}} (\eta(h - \ell) - \eta(h - 10 \log R - \ell)) (\phi_{\leq \ell-10} \phi_\ell^\dagger - \phi_\ell \phi_{\leq \ell-10}^\dagger).$$

This results in the new equation

$$\square \left(U_{<k}^{(\phi^n, <K_*)} \overline{\phi_k^{n, <K_*, R, M}} \right) = U_{<k}^{(\phi^n, <K_*, R)} F_k^{(R, M)} (\phi^{n, <K_*}, \nabla_{t,x} \phi^{n, <K_*}).$$

In order to pass to an equation for $\eta^{Kn, <K_*}$, we need to subtract off the bulk terms. This is straightforward for the low frequency term u^n as well as the asymptotic terms

$$\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right),$$

but more delicate for the bulk term $\epsilon_{bounded}^{<K_*}$. Here we make the following

Observation 1: For any fixed k and n there exists a (constant) matrix $G_k^{(n)} \in SO(m+1)$ with the property that as $n \rightarrow \infty$

$$\sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_k \| P_{k+a} (U_{<k}^{(\phi^n, <K_*)} \overline{\epsilon_{bounded,k}^{<K_*, R, M}} - G_k^{(n)} \cdot U_{<k}^{(\epsilon_{bounded}^{<K_*}, R)} \overline{\epsilon_{bounded,k}^{<K_*, R, M}}) \|_{S_{k+a}([0, R])}^2 \right)^{\frac{1}{2}} \\ = o(1) + c_R \| \eta^{Kn, <K_*} \|_{S([0, R] \times \mathbb{R}^2)} + \delta_4 e^{-R^{1000} - 2K_*}$$

for suitable $\sigma > 0$. Also, there is a partition of $[0, R]$ into $O(R^{100} + K_*^{100})$ many time intervals I_j such that we have for all j

$$\left(\sum_k \| U_{<k}^{(\phi^n, <K_*)} F_k^{(R, M)} (\epsilon_{bounded}^{<K_*}, \nabla \epsilon_{bounded}^{<K_*}) - G_k^{(n)} \cdot U_{<k}^{(\epsilon_{bounded}^{<K_*}, R)} F_k^{(R, M)} (\epsilon_{bounded}^{<K_*}, \nabla \epsilon_{bounded}^{<K_*}) \|_{L_t^1 L_x^2[I_j]}^2 \right)^{\frac{1}{2}} \\ = o(1) + c_R \| \eta^{Kn, <K_*} \|_S + \delta_4 e^{-R^{1000} - 2K_*},$$

where the error $o(1)$ is in the sense as $n \rightarrow \infty$. Also, c_R is a coefficient with $\lim_{R \rightarrow \infty} c_R = 0$. In fact, the matrix $G_k^{(n)}$ does not depend on M but it may implicitly depend on R, K_* .

We observe here that the factors $G_k^{(n)}$ are analogous to the exponential corrections $e^{i\gamma \delta_2 nab}$ in Proposition 9.24 in [17]. We complement this with the following observation of a similar flavor:

Observation 2: We have the asymptotic relations as $n \rightarrow \infty$

$$\sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_k \| P_{k+a} (U_{<k}^{(\phi^n, <K_*)} \overline{u_k^{n, R, M}} - U_{<k}^{(u^n, R)} \overline{u_k^{n, R, M}}) \|_S^2 \right)^{\frac{1}{2}} = o(1), \\ \sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_k \| P_{k+a} (U_{<k}^{(\phi^n, <K_*)} F_k^{(R, M)} (u^n, \nabla u^n) - U_{<k}^{(u^n, R)} F_k^{(R, M)} (u^n, \nabla u^n)) \|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}} = o(1),$$

where $\overline{u_k^{n, R, M}}$ is defined in analogy to $\overline{\phi_k^{n, <K_*, R, M}}$.

We leave the proofs of these observations for later. Then the proof of Lemma 7.21 shall follow from the following steps:

Step 1 of proof of Lemma 7.21: Fix a time slice $[0, R] \times \mathbb{R}^2$. Given $\delta_4 > 0$ (which we will later choose sufficiently small depending on $\delta_3, \|\epsilon_{\text{bounded}}^{<K_*}\|_S, \|u^n\|_S$), then we have for K sufficiently large that

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \left(\sum_k \left\| \square(U_{<k}^{(\phi^n, <K_*, R)} \overline{\phi_k^{n, <K_*, R, M}} - U_{<k}^{(u^n, R)} \overline{u_k^{n, R, M}} - G_k^{(n)} \cdot U_{<k}^{(\epsilon^{<K_*}, R)} \overline{\epsilon_{\text{bounded}, k}^{<K_*, R, M}}) \right\|_{L_t^1 L_x^2([0, R])}^2 \right)^{\frac{1}{2}} \\ & \leq \delta_4 e^{-R^{1000} - 2K_*} C_{11}(\|\epsilon_{\text{bounded}}^{<K_*}\|_S, \|u^n\|_S) + R^{41} (\log R + K_*)^{\frac{1}{2}} \|\eta^{Kn, <K_*}\|_{S([0, R])} C_{10}(\|\epsilon_{\text{bounded}}^{<K_*}\|_S, \|u^n\|_S) \\ & \quad + D(\|\epsilon_{\text{bounded}}^{<K_*}\|_S, \|u^n\|_S) \|\eta^{Kn, <K_*}\|_{S([0, R])}^3. \end{aligned}$$

To see this, we start by observing the identity

$$\square(G_k^{(n)} \cdot U_{<k}^{(\epsilon_{\text{bounded}}^{<K_*}, R)} \overline{\epsilon_{\text{bounded}, k}^{<K_*, R, M}}) = G_k^{(n)} \cdot U_{<k}^{(\epsilon_{\text{bounded}}^{<K_*}, R)} F_k^{(R, M)}(\epsilon_{\text{bounded}}^{<K_*}, \nabla \epsilon_{\text{bounded}}^{<K_*}),$$

and so

$$\begin{aligned} & \square(U_{<k}^{(\phi^n, <K_*, R)} \overline{\phi_k^{n, <K_*, R, M}} - U_{<k}^{(u^n, R)} \overline{u_k^{n, R, M}} - G_k^{(n)} \cdot U_{<k}^{(\epsilon_{\text{bounded}}^{<K_*}, R)} \overline{\epsilon_{\text{bounded}, k}^{<K_*, R, M}}) \\ & = U_{<k}^{(\phi^n, <K_*, R)} F_k^{(R, M)}(\phi^{n, <K_*}, \nabla \phi^{n, <K_*}) - G_k^{(n)} \cdot U_{<k}^{(\epsilon_{\text{bounded}}^{<K_*}, R)} F_k^{(R, M)}(\epsilon_{\text{bounded}}^{<K_*}, \nabla \epsilon_{\text{bounded}}^{<K_*}) - U_{<k}^{(u^n, R)} F_k^{(R, M)}(u^n, \nabla u^n) \\ & = U_{<k}^{(\phi^n, <K_*, R)} (F_k^{(R, M)}(\phi^{n, <K_*}, \nabla \phi^{n, <K_*}) - F_k^{(R, M)}(\epsilon_{\text{bounded}}^{<K_*}, \nabla \epsilon_{\text{bounded}}^{<K_*}) - F_k^{(R, M)}(u^n, \nabla u^n)) + A_k^{(R, M, n)}, \end{aligned}$$

where Observations 1 and 2 imply that after restriction to any one of $O(R^{100} + K_*^{100})$ many time intervals I_j we get

$$\lim_{R, n \rightarrow \infty} \left(\sum_k \|A_k^{(R, M, n)}\|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}} = 0.$$

This time restriction step will be carried out explicitly in *Step 5* below.

Thus, it now remains to bound the term

$$(7.47) \quad U_{<k}^{(\phi^n, <K_*, R)} (F_k^{(R, M)}(\phi^{n, <K_*}, \nabla \phi^{n, <K_*}) - F_k^{(R, M)}(\epsilon_{\text{bounded}}^{<K_*}, \nabla \epsilon_{\text{bounded}}^{<K_*}) - F_k^{(R, M)}(u^n, \nabla u^n)).$$

Here we distinguish between different interaction terms, first amongst the bulk terms in the decomposition (7.39). Before doing so, we quickly comment on the fact that the nonlinearity $F_k^{(R, M)}(\phi^{n, <K_*}, \nabla \phi^{n, <K_*})$ now involves some additional terms on account of the fact that we modified $\overline{\phi^{n, <K_*}}$ to $\overline{\phi^{n, <K_*, R, M}}$ as well as $U_{<k}^{(\phi^n, <K_*)}$ to $U_{<k}^{(\phi^n, <K_*, R)}$. For the former modification these terms are schematically of the form

$$(7.48) \quad \begin{aligned} & P_k(\phi_{[k-10 \log R, k]} \sum_{k_1=k_2+O(1) \in [k, k+\log R]} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_2}), \quad \sum_{k_1 \in [k-10 \log M, k]} P_k(\chi_{\{r \leq 2^{-k} M\}} \phi_{<k_1-10} \nabla \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_k) \\ & P_k(\phi_{<k} \sum_{k < k_1=k_2+O(1) < k+\log M} \chi_{\{r \leq 2^{-k} M\}} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_2}), \quad P_k(\phi_{<k} \sum_{k < k_1=k_2+O(1) < k+\log M} \chi'_{\{r \leq 2^{-k} M\}} \phi_{k_1}^\dagger \nabla_{t,x} \phi_{k_2}) \\ & P_k(\sum_{k_1 \in [k-10 \log M, k]} \chi'_{\{r \leq 2^{-k} M\}} \phi_{<k_1-10} \nabla \phi_{k_1}^\dagger \partial^\alpha \phi_k) \end{aligned}$$

as well as additional terms where χ' is replaced by χ'' at the expense of the operator $\nabla_{t,x}$. For the latter modification we get additional terms of the schematic form

$$\begin{aligned} & \sum_{k_1 \in [k-10 \log R, k]} \phi_{<k_1-10} \chi_{\{r \leq 2^{-k_1} R^{40}\}} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_k, \quad \sum_{k_1 \in [k-10 \log R, k]} \phi_{<k_1-10} \chi'_{\{r \leq 2^{-k_1} R^{40}\}} \phi_{k_1}^\dagger \nabla_{t,x} \phi_k \\ & \sum_{k_1 \in [k-10 \log R, k]} \phi_{<k_1-10} \chi''_{\{r \leq 2^{-k_1} R^{40}\}} \phi_{k_1}^\dagger \phi_k. \end{aligned}$$

We claim that all of these terms admit good bounds in terms of the frequency square-summed $L_t^1 L_x^2$ -norm, albeit possibly at the expense of a power of R or M , and that they all have the important divisibility

property. By this of course we mean that if one of the factors ϕ_{k_j} is replaced by η_{k_j} , then we can divide time into $O(M^K + R^K)$ many intervals for some explicit K such that the corresponding square sum $L_t^1 L_x^2$ -norm restricted to such a time slice is bounded by $\ll \|\eta\|_S$. The implicit constant in $O(M^K + R^K)$ depends only on $\|\phi\|_S$. We quickly explain this for the first term in (7.48), the others being handled similarly. Thus write

$$\begin{aligned}
(7.49) \quad & P_k(\phi_{[k-10\log R, k]} \sum_{k_1=k_2+O(1)\in[k, k+\log R]} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_2}) \\
&= P_k(\chi_{\{r \leq R^{40} 2^{-k}\}} \phi_{[k-10\log R, k]} \sum_{k_1=k_2+O(1)\in[k, k+\log R]} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_2}) \\
&\quad + P_k(\chi_{\{r > R^{40} 2^{-k}\}} \phi_{[k-10\log R, k]} \sum_{k_1=k_2+O(1)\in[k, k+\log R]} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_2}).
\end{aligned}$$

Then to estimate the second term on the right-hand side we use the null form estimate (4.1) to infer that

$$\begin{aligned}
& \|P_k(\chi_{\{r > R^{40} 2^{-k}\}} \phi_{[k-10\log R, k]} \sum_{k_1=k_2+O(1)\in[k, k+\log R]} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_2})\|_{L_t^1 L_x^2} \\
& \lesssim 2^{\frac{k}{3}} \|\chi_{\{r > R^{40} 2^{-k}\}} \phi_{[k-10\log R, k]}\|_{L_t^3 L_x^\infty} \left\| \sum_{k_1=k_2+O(1)\in[k, k+\log R]} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_2} \right\|_{L_{t,x}^{\frac{3}{2}}} \\
& \lesssim R^{-1} \left(\sum_{\ell \in [k-10\log R, k]} \|\phi_\ell\|_{S_\ell}^2 \right)^{\frac{1}{2}} \sum_{k_1} \|\phi_{k_1}\|_{S_{k_1}}^2.
\end{aligned}$$

Here we have used an interpolate between the improved L_x^∞ -bound

$$\|\chi_{\{r \geq 2^{-k} R\}} P_k \phi\|_{L_{t,x}^\infty} \lesssim R^{-\frac{1}{2}} \|P_k \phi\|_{\dot{H}_x^1}$$

and the $L_t^{2+} L_x^\infty$ Strichartz estimate to get improved control over $\|\chi_{\{r \geq 2^{-k} R\}} P_k \phi\|_{L_t^3 L_x^\infty}$. Finally, square-summing over $k \in \mathbb{Z}$, we get

$$\left(\sum_{k \in \mathbb{Z}} \|P_k(\chi_{\{r > R^{40} 2^{-k}\}} \phi_{[k-10\log R, k]} \sum_{k_1=k_2+O(1)\in[k, k+\log R]} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_2})\|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}} \lesssim R^{-1} \log(R) \|\phi\|_S^3,$$

which of course gains smallness for $R \gg 1$. To bound the first term on the right-hand side of (7.49) further above, we use the local energy decay norm

$$\begin{aligned}
& \|P_k(\chi_{\{r \leq R^{40} 2^{-k}\}} \phi_{[k-10\log R, k]} \sum_{k_1=k_2+O(1)\in[k, k+\log R]} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_2})\|_{L_t^1 L_x^2} \\
& \lesssim R^{40} \|\phi_{[k-10\log R, k]}\|_{L_{t,x}^\infty} \left(\sum_{k+\log R > k_1 > k} \sum_{\ell < 40\log R - k} (2^{\frac{k_1+\ell}{2}} R^{-20} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_{k_1}\|_{L_{t,x}^2})^2 \right) \\
& \lesssim R^{40} \|\phi_{[k-10\log R, k]}\|_{L_{t,x}^\infty} \left(\sum_{k+\log R > k_1 > k} \sum_{\ell < 40\log R - k_1} (2^{\frac{k_1+\ell}{2}} R^{-20} \|\chi_{\{r \sim 2^\ell\}} \nabla_{t,x} \phi_{k_1}\|_{L_{t,x}^2})^2 \right).
\end{aligned}$$

Then square summing over $k \in \mathbb{Z}$ results in the bound

$$\left(\sum_{k \in \mathbb{Z}} \|P_k(\chi_{\{r \leq R^{40} 2^{-k}\}} \phi_{[k-10\log R, k]} \sum_{k_1=k_2+O(1)>k} \partial_\alpha \phi_{k_1}^\dagger \partial^\alpha \phi_{k_2})\|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}} \lesssim R^{40} \log(R) \|\phi\|_S^3.$$

Moreover, the quantity

$$\sum_{k_1} \sum_{\ell < 40\log R - k_1} \left(2^{\frac{k_1+\ell}{2}} R^{-20} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_{k_1}\|_{L_{t,x}^2} \right)^2 \lesssim \sum_{k_1} \sup_{\ell \in \mathbb{Z}} \|\chi_{\{r \sim 2^\ell\}} r^{-\frac{1}{2}} \nabla_{t,x} \phi_{k_1}\|_{L_{t,x}^2}^2 \lesssim \|\phi\|_S^2$$

is easily seen to have the divisibility property.

Back to controlling (7.47), we now discuss how to bound the various interactions:

(i) *Interactions between u^n and $\epsilon_{\text{bounded}}^{<K_*}$.* These are either terms in which u^n appears in one of the lowest frequency “non-perturbative” factors such as

$$P_k \left((u_{<k-10}^n - p_*) \sum_{k_{1,2} > k+10} P_{k_1} (P_{<k_1-10} \epsilon_{\text{bounded}}^{<K_*} \sum_{k_4=k_5+O(1)} (\partial_\alpha \epsilon_{\text{bounded},k_4}^{<K_*})^\dagger \partial^\alpha \epsilon_{\text{bounded},k_5}^{<K_*})^\dagger P_{k_2} \epsilon_{\text{bounded}}^{<K_*} \right),$$

where one can only place $(u_{<k-10}^n - p_*)$ into $L_{t,x}^\infty$ or else they are terms where u^n appears in a “perturbative factor” such as

$$P_k \left(\epsilon_{\text{bounded}, <k-10}^{<K_*} \sum_{k_{1,2} > k+10} P_{k_1} (P_{<k_1-10} \epsilon_{\text{bounded}}^{<K_*} \sum_{k_4=k_5+O(1)} \partial_\alpha (u_{k_4}^n)^\dagger \partial^\alpha u_{k_5}^n)^\dagger P_{k_2} u_{<k-10}^n \right).$$

For the former type of term one gets smallness for the square-summed $L_t^1 L_x^2$ norm from the smallness of $u_{<k-10}^n - p_*$ on the support of $P_k \epsilon_{\text{bounded}}^{<K_*}$ for $k \in \mathbb{Z}$ of moderate size, while for extremely small or large k one exploits the smallness of $P_k \epsilon_{\text{bounded}}^{<K_*}$. For the second type of term, one may assume that the frequencies $2^{k_2}, 2^{k_4}, 2^{k_5}$ applied to u^n are extremely small, in which case one gains by placing the nonlinearity into $L_t^\infty L_x^2$ instead of $L_t^1 L_x^2$ and making up by multiplying with R (which is held fixed) and then letting $n \rightarrow \infty$.

(ii) *Interactions between u^n and $\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right)$.* Here, if u^n appears in a “non-perturbative factor” and not in a “perturbative factor”, then we exploit the pointwise decay of the expression $\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right)$. In fact, observe that due to the essential frequency localization (up to exponentially decaying tails) we infer

$$\lim_{n \rightarrow \infty} \left\| \sum_k (U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} \right) \right\|_{L_{t,x}^\infty([0,R] \times \mathbb{R}^2)} = 0.$$

The same holds for the remainder term $\sum_k (U_{<k}^{(\phi^n, <K_*)})^\dagger \tilde{\epsilon}_{B,k}^{Kn, <K_*}$ since $\tilde{\epsilon}_{B,k}^{Kn, <K_*}(0, \cdot)$ converges weakly to zero for all $k \in \mathbb{Z}$. To close the case in which all “perturbative factors” are of the form

$$\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right),$$

we achieve $L_t^1 L_x^2$ for the product by placing it into $L_t^\infty L_x^2$ and multiplying with R , and the former norm can be bounded by placing all factors into $L_t^\infty L_x^6$ (for cubic terms, and similarly for higher order terms) by interpolating between $L_t^\infty L_x^2$ and $L_{t,x}^\infty$, with the latter norm being asymptotically vanishing (as $n \rightarrow \infty$) due to the preceding remarks.

If u^n appears in at least one “perturbative factor” while $\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right)$ appears in the “non-perturbative factor”, we close again by using the asymptotic vanishing of this term in $L_{t,x}^\infty$ if not all “perturbative factors” are u^n . If all “perturbative factors” are frequency localized versions of u^n , we gain smallness on account of the frequency localization of

$$\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right)$$

to frequencies $\geq \mu_n$ up to exponentially decaying errors.

(iii) *Interactions between $\epsilon_{\text{bounded}}^{<K_*}$ and $\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right)$ as well as pure self-interactions of $\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right)$.* These are handled exactly like the preceding case.

We emphasize that the discussion up to this point applies to all terms occurring inside $F_k^{(R,M)}$, also to the additional terms stemming from modifying F_k to $F_k^{(R,M)}$, in light of the discussion preceding (i).

(iv) *Interactions involving $\eta^{Kn, < K_*}$.* These are the most delicate ones, but we can again get rid of most terms:

(iv.1) *Interactions involving $\eta^{Kn, < K_*}$ and at least one factor u^n .* Assume that the factor u^n is in a non-perturbative position, and all other perturbative factors are occupied by (frequency localizations of) $\eta^{Kn, < K_*}$. Write this term schematically as $(u_{< k-10}^n - p_*)F_k^{(R, M)}(\eta^{Kn, < K_*})$. Then we get

$$\left(\sum_k \left\| (u_{< k-10}^n - p_*)F_k^{(R, M)}(\eta^{Kn, < K_*}) \right\|_{L_t^1 L_x^2([0, R])}^2 \right)^{\frac{1}{2}} \lesssim \left\| \eta^{Kn, < K_*} \right\|_S^3.$$

On the other hand, if at least one of the other perturbative factors is occupied by a $\epsilon_{bounded}^{< K_*}$ or an asymptotically vanishing term $\sum_k \tilde{P}_k \left((U_{< k}^{(\phi^n, < K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b, k}^{Kn, < K_*} + \tilde{\epsilon}_{B, k}^{Kn, < K_*} \right) \right)$, we estimate it by

$$\begin{aligned} & \left(\sum_k \left\| u_{< k-10}^n F_k^{(R, M)}(\eta^{Kn, < K_*}, \epsilon_{bounded}^{< K_*}, \sum_k (U_{< k}^{(\phi^n, < K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b, k}^{Kn, < K_*} + \tilde{\epsilon}_{B, k}^{Kn, < K_*} \right) \right\|_{L_t^1 L_x^2([0, R])}^2 \right)^{\frac{1}{2}} \\ & \leq C_1 \left\| \eta^{Kn, < K_*} \right\|_S^2 + C_2 \left\| \eta^{Kn, < K_*} \right\|_S, \end{aligned}$$

where $C_{1,2}$ depend only on $\left\| \epsilon_{bounded}^{< K_*} \right\|_S$ provided n is sufficiently large to kill the effect of the asymptotically vanishing terms.

Next, consider the terms where at least one perturbative factor is occupied by u^n and where there is at least one factor $\eta^{Kn, < K_*}$. If this latter factor is in a perturbative position, then we bound the term by

$$\lesssim C_3 \left(\|u^n\|_S, \left\| \epsilon_{bounded}^{< K_*} \right\|_S \right) \left\| \eta^{Kn, < K_*} \right\|_S^2 + C_4 \left(\|u^n\|_S, \left\| \epsilon_{bounded}^{< K_*} \right\|_S \right) \left\| \eta^{Kn, < K_*} \right\|_S.$$

On the other hand, if the factor $\eta^{Kn, < K_*}$ is in a non perturbative position and all perturbative factors are occupied by u^n , then we have to complement this argument with a pure energy type bound just as in the proof of the bound for the lowest frequency non-atomic part in Proposition 7.9. We provide this in the next item:

(iv.2) *Choosing K and n large enough, we have for all $k < \mu_n$ that*

$$\begin{aligned} & \left\| P_k \eta^{Kn, < K_*} \right\|_{L_t^\infty L_x^2([0, R])} \\ & \ll \delta_4 e^{-R^{1000} - 2K_*} 2^{-\sigma|k-b_L|} + R 2^k \left(C_5 \left(\|u^n\|_S, \left\| \epsilon_{bounded}^{< K_*} \right\|_S \right) \left\| \eta^{nK} \right\|_S + C_6 \left(\|u^n\|_S, \left\| \epsilon_{bounded}^{< K_*} \right\|_S \right) \left\| \eta^{nK} \right\|_S^3 \right) \\ & \ll \delta_4 e^{-R^{1000} - 2K_*} 2^{-\sigma|k-b_L|} \end{aligned}$$

for a suitable absolute constant $\sigma > 0$. Here δ_4 will be determined in terms of δ_3 as well as $\|u^n\|_S$. In fact, this is much simpler than the proof of the corresponding bound for the lowest frequency non-atomic part in Lemma 7.10. Replacing ϵ by $\eta^{Kn, < K_*}$ there and ϕ by

$$u = u^n - p_* + \epsilon_{bounded}^{< K_*} + \sum_k (U_{< k}^{(\phi^n, < K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b, k}^{Kn, < K_*} + \tilde{\epsilon}_{B, k}^{Kn, < K_*} \right),$$

then fixing a very low frequency $k < \mu_n$ we have to bound integrals of the form

$$\int_0^R \int_{\mathbb{R}^2} P_k(\eta^{Kn, < K_*} \partial_\alpha u^\dagger \partial^\alpha u)^\dagger \partial_t P_k \eta^{Kn, < K_*} dx dt$$

as well as other, similar ones. But the preceding expression is bounded by

$$\begin{aligned} & \left| \int_0^R \int_{\mathbb{R}^2} P_k(\eta^{Kn, < K_*} \partial_\alpha u^\dagger \partial^\alpha u)^\dagger \partial_t P_k \eta^{Kn, < K_*} dx dt \right| \\ & \leq R \left\| \eta^{Kn, < K_*} \right\|_{L_{t,x}^\infty([0, R])} \left\| \partial_\alpha u \partial^\alpha u \right\|_{L_t^\infty L_x^{1+}([0, R])} \left\| \partial_t P_k \eta^{Kn, < K_*} \right\|_{L_t^\infty L_x^{\infty-}([0, R])} \end{aligned}$$

and this in turn is bounded by the expression further above. Indeed, observe that the terms

$$\sum_k (U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right)$$

asymptotically do not contribute, on account of their asymptotic vanishing property. The term $\delta_4 e^{-R^{1000}-2K_*} 2^{-\sigma|k-b_L|}$ comes from the initial data $\eta^{Kn,<K_*}[0]$ at time $t = 0$, which can be made small by picking K large enough on account of Lemma 7.14. The other integrals in the energy identity are handled similarly.

We can now close case (iv.1):

(iv.3) *Conclusion of case (iv.1).* Consider now a term of the schematic form $\eta_{<k-10}^{Kn,<K_*} F_k^{(R,M)}(u^n, \nabla u^n)$. We bound this by

$$\begin{aligned} \left(\sum_k \|\eta_{<k-10}^{Kn,<K_*} F_k^{(R,M)}(u^n, \nabla u^n)\|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}} &\lesssim \left(\sum_{k \geq \mu_n} \|\eta_{<k-10}^{Kn,<K_*} F_k^{(R,M)}(u^n, \nabla u^n)\|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}} \\ &\quad + \left(\sum_{k < \mu_n} \|\eta_{<k-10}^{Kn,<K_*} F_k^{(R,M)}(u^n, \nabla u^n)\|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Then for the first term on the right we have

$$\lim_{n \rightarrow \infty} \left(\sum_{k \geq \mu_n} \|\eta_{<k-10}^{Kn,<K_*} F_k^{(R,M)}(u^n, \nabla u^n)\|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}} = 0,$$

while for the second term on the right, we use (iv.2) to bound it by

$$\limsup_{n \rightarrow \infty} \left(\sum_{k < \mu_n} \|\eta_{<k-10}^{Kn,<K_*} F_k^{(R,M)}(u^n, \nabla u^n)\|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}} \lesssim \delta_4 e^{-R^{1000}-2K_*} \|u^n\|_S^3.$$

(v) *Interactions between $\eta^{Kn,<K_*}$ and $\epsilon_{bounded}^{<K_*}$.* This is the most delicate case on account of the situations where $\eta^{Kn,<K_*}$ is in a non-perturbative position while the perturbative factors are all $\epsilon_{bounded}^{<K_*}$. In fact, it is in this situation where the very precise information about $\eta^{Kn,<K_*}$ at $t = 0$ from Lemma 7.14 becomes really important. The basic idea is to partition the time slice $[0, R] \times \mathbb{R}^2$ into a cylinder $[0, R] \times B_{R_*}$ for some very large R_* chosen such that $\epsilon_{bounded}^{<K_*}$ is small in the complement of B_{R_*} , and the set $[0, R] \times B_{R_*}^c$. More specifically, R_* will be chosen depending on R and δ_3 , and R will then later be chosen to control the solution on all of $[0, t_1^n - T_*]$.

Now given R and $\delta_3 > 0$, pick R_* sufficiently large such that

$$\left(\sum_k \|\chi_{B_{R_*}^c} F_k^{(R,M)}(\epsilon_{bounded}^{<K_*})\|_{L_t^1 L_x^2([0,R] \times \mathbb{R}^2)}^2 \right)^{\frac{1}{2}} \ll \delta_4 e^{-R^{1000}-2K_*},$$

where $\eta_{<k-10}^{Kn,<K_*} F_k^{(R,M)}(\epsilon_{bounded}^{<K_*})$ is a schematically written term in which $\eta_{<k-10}^{Kn,<K_*}$ is a non-perturbative factor while all the perturbative positions are occupied by $\epsilon_{bounded}^{<K_*}$. Then due to the simple $L_{t,x}^\infty$ -bound $\|\eta^{Kn,<K_*}\|_{L_{t,x}^\infty} \lesssim 1$ provided K, n are sufficiently large, we get the bound

$$\left(\sum_k \|\chi_{B_{R_*}^c} \eta_{<k-10}^{Kn,<K_*} F_k^{(R,M)}(\epsilon_{bounded}^{<K_*})\|_{L_t^1 L_x^2([0,R] \times \mathbb{R}^2)}^2 \right)^{\frac{1}{2}} \ll \delta_4 e^{-R^{1000}-2K_*}.$$

Thus, consider now the region $[0, R] \times B_{R_*}$ where $\eta^{Kn,<K_*}$ and $\epsilon_{bounded}^{<K_*}$ interact strongly. Then the smallness will have to come from the factor $\eta_{<k-10}^{Kn,<K_*}$. In fact, use the splitting

$$\chi_{B_{R_*}} \eta_{<k-10}^{Kn,<K_*} F_k^{(R,M)}(\epsilon_{bounded}^{<K_*}) = \chi_{B_{R_*}} \eta_{[-10 \log R, k-10]}^{Kn,<K_*} F_k^{(R,M)}(\epsilon_{bounded}^{<K_*}) + \chi_{B_{R_*}} \eta_{<-10 \log R}^{Kn,<K_*} F_k^{(R,M)}(\epsilon_{bounded}^{<K_*}).$$

Then we control the first term on the right with the estimate

$$\begin{aligned}
& \left(\sum_k \left\| \chi_{B_{R_*}} \eta_{[-10 \log R, k-10]}^{Kn, < K_*} F_k^{(R, M)}(\epsilon_{bounded}^{< K_*}) \right\|_{L_t^1 L_x^2([0, R] \times \mathbb{R}^2)}^2 \right)^{\frac{1}{2}} \\
& \lesssim R^{41} (\log R + K_*)^{\frac{1}{2}} \left\| \eta^{Kn, < K_*} \right\|_S C_7 \left(\left\| \epsilon_{bounded}^{< K_*} \right\|_S \right) \\
& \quad + \left(\sum_{k > K_*} \left\| \chi_{B_{R_*}} \eta_{[-10 \log R, k-10]}^{Kn, < K_*} F_k^{(R, M)}(\epsilon_{bounded}^{< K_*}) \right\|_{L_t^1 L_x^2([0, R] \times \mathbb{R}^2)}^2 \right)^{\frac{1}{2}} \\
& \lesssim R^{41} (\log R + K_*)^{\frac{1}{2}} \left\| \eta^{Kn, < K_*} \right\|_S C_8 \left(\left\| \epsilon_{bounded}^{< K_*} \right\|_S \right)
\end{aligned}$$

on account of the fact that the norms $\left\| F_k^{(R, M)}(\epsilon_{bounded}^{< K_*}) \right\|_{L_t^1 L_x^2([0, R] \times \mathbb{R}^2)}$ are exponentially decreasing for $k \geq K_*$ by our enforced frequency localization. For the second term on the right we use that for $t \in [0, R]$

$$\left| \chi_{B_{R_*}} \eta_{< -10 \log R}^{Kn, < K_*}(t, \cdot) \right| \leq R \left\| \partial_t \eta_{< -10 \log R}^{Kn, < K_*} \right\|_{L_{t,x}^\infty([0, R] \times \mathbb{R}^2)} + \left\| \chi_{B_{R_*}} \eta_{< -10 \log R}^{Kn, < K_*}(0, \cdot) \right\|_{L_x^\infty}.$$

Since we may choose $R_* \gg R^{10}$ we may essentially move the cutoff $\chi_{B_{R_*}}$ past the frequency localizer (up to exponentially decaying tails) and so picking K and R_* large enough we may assume that

$$\left\| \chi_{B_{R_*}} \eta_{< -10 \log R}^{Kn, < K_*}(0, \cdot) \right\|_{L_x^\infty} \ll \delta_4 e^{-R^{1000} - 2K_*}.$$

The first term on the right on the other hand may be bounded by

$$R \left\| \partial_t \eta_{< -10 \log R}^{Kn, < K_*} \right\|_{L_{t,x}^\infty([0, R] \times \mathbb{R}^2)} \lesssim R^{-9} \left\| \eta^{Kn, < K_*} \right\|_S,$$

and so we obtain

$$\left(\sum_k \left\| \chi_{B_{R_*}} \eta_{< -10 \log R}^{Kn, < K_*} F_k^{(R, M)}(\epsilon_{bounded}^{< K_*}) \right\|_{L_t^1 L_x^2([0, R] \times \mathbb{R}^2)}^2 \right)^{\frac{1}{2}} \lesssim \left(\delta_4 e^{-R^{1000} - 2K_*} + R^{-9} \left\| \eta^{Kn, < K_*} \right\|_S \right) C_9 \left(\left\| \epsilon_{bounded}^{< K_*} \right\|_S \right).$$

In total we have obtained the following bound

$$\begin{aligned}
& \left(\sum_k \left\| \eta_{< -10 \log R}^{Kn, < K_*} F_k^{(R, M)}(\epsilon_{bounded}^{< K_*}) \right\|_{L_t^1 L_x^2([0, R] \times \mathbb{R}^2)}^2 \right)^{\frac{1}{2}} \\
& \lesssim \delta_4 e^{-R^{1000} - 2K_*} C_9 \left(\left\| \epsilon_{bounded}^{< K_*} \right\|_S \right) + R^{41} (\log R + K_*)^{\frac{1}{2}} \left\| \eta^{Kn, < K_*} \right\|_S C_{10} \left(\left\| \epsilon_{bounded}^{< K_*} \right\|_S \right).
\end{aligned}$$

Combining the cases (i) – (v) and the remarks preceding them, we finally obtain the bound

$$\begin{aligned}
(7.50) \quad & \left(\sum_k \left\| \square(U_{< k}^{(\phi^{n, < K_*}, R)} \overline{\phi_k^{n, < K_*, M, R}} - U_{< k}^{(u^n, R)} \overline{u_k^{n, M, R}} - G_k^{(n)} \cdot U_{< k}^{(\epsilon^{< K_*}, R)} \overline{\epsilon_{bounded, k}^{< K_*, M, R}}) \right\|_{L_t^1 L_x^2([0, R])}^2 \right)^{\frac{1}{2}} \\
& \lesssim \delta_4 e^{-R^{1000} - 2K_*} C_{11} \left(\left\| \epsilon_{bounded}^{< K_*} \right\|_S, \|u^n\|_S \right) + R^{41} (\log R + K_*)^{\frac{1}{2}} \left\| \eta^{Kn, < K_*} \right\|_S C_{10} \left(\left\| \epsilon_{bounded}^{< K_*} \right\|_S, \|u^n\|_S \right) \\
& \quad + D \left(\left\| \epsilon_{bounded}^{< K_*} \right\|_S, \|u^n\|_S \right) \left\| \eta^{Kn, < K_*} \right\|_S^3.
\end{aligned}$$

The bound established thus far in Step 1 of the proof of Lemma 7.21 is not quite enough yet due to the fact that the terms $U_{< k}^{(\phi^{n, < K_*}, R)} \overline{\phi_k^{n, < K_*, M, R}}$ etc. are not sharply frequency localized on account of the lack of

frequency localizations of the gauge transformations $U_{<k}^{(\phi^n, <K_*, R)}$ etc. However, a straightforward Littlewood-Paley trichotomy reveals that we get the somewhat sharper estimate

$$(7.51) \quad \begin{aligned} & \sum_a 2^{\sigma|a|} \left(\sum_k \left\| P_{k+a} \square \left(U_{<k}^{(\phi^n, <K_*, R)} \overline{\phi_k^{n, <K_*, M, R}} - U_{<k}^{(u^n, R)} \overline{u_k^{n, M, R}} - G_k^{(n)} \cdot U_{<k}^{(\epsilon^{<K_*}, R)} \overline{\epsilon_{bounded, k}^{<K_*, M, R}} \right) \right\|_{L_t^1 L_x^2([0, R])}^2 \right)^{\frac{1}{2}} \\ & \lesssim \delta_4 e^{-R^{1000} - 2K_*} C_{11} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) + R^{41} (\log R + K_*)^{\frac{1}{2}} \|\eta^{Kn, <K_*}\|_S C_{10} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) \\ & \quad + D \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) \|\eta^{Kn, <K_*}\|_S^3 \end{aligned}$$

for any $\sigma < 1$.

Step 2 of the proof of Lemma 7.21: *The preceding step implies via the basic energy estimate (3.1) that for M, K, n large enough the following bound holds*

$$(7.52) \quad \begin{aligned} & \sum_a 2^{\sigma|a|} \left(\sum_k \left\| P_{k+a} \left(U_{<k}^{(\phi^n, <K_*, R)} \left((\eta_k^{Kn, <K_*} + \epsilon_{bounded, k}^{<K_*})^{M, R} - \overline{\epsilon_{bounded, k}^{<K_*, M, R}} \right) \right) \right\|_{S([0, R])}^2 \right)^{\frac{1}{2}} \\ & \lesssim \delta_4 e^{-R^{1000} - 2K_*} C_{11} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) + R^{41} (\log R + K_*)^{\frac{1}{2}} \|\eta^{Kn, <K_*}\|_{S([0, R])} C_{10} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) \\ & \quad + E \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) \|\eta^{Kn, <K_*}\|_{S([0, R])}^3 + c_R \|\eta^{Kn, <K_*}\|_S, \end{aligned}$$

where

$$\lim_{R \rightarrow \infty} c_R = 0.$$

This follows essentially from the basic energy inequality (3.1). In fact, the self-interactions of $\epsilon_{bounded}^{<K_*}$ inside $\overline{\phi_k^{n, <K_*, M, R}}$ are cancelled by $G_k^{(n)} \cdot U_{<k}^{(\epsilon^{<K_*}, R)} \overline{\epsilon_{bounded, k}^{<K_*, M, R}}$ via *Observation 1*.

The self-interactions of u^n inside $\overline{\phi_k^{n, <K_*, M, R}}$ are cancelled by $U_{<k}^{(u^n, R)} \overline{u_k^{n, M, R}}$ via *Observation 2*. Next, for the interactions between $\epsilon_{bounded}^{<K_*}$ and $u^{n, M}$ inside $\overline{\phi_k^{n, <K_*, M, R}}$ we use the smallness of $u^n - p_*$ on the support of $\epsilon_{bounded}^{<K_*}$ (i.e. on bounded sets $\{r = O(1)\}$) provided all perturbative factors are occupied by $\epsilon_{bounded}^{<K_*}$, while in case that all perturbative factors are occupied by u^n we gain smallness by exploiting the frequency localization of $\epsilon_{bounded}^{<K_*}$ to the range $k \geq \mu_n$ up to exponential tails. In case that there are perturbative factors occupied by both u^n as well $\epsilon_{bounded}^{<K_*}$, the smallness also follows from the essentially disjoint frequency supports (up to exponentially decaying tails) of these functions.

Furthermore all interactions of $u^{n, M}$, $\epsilon_{bounded}^{<K_*}$ or $\eta^{Kn, <K_*}$ with $\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b, k}^{Kn, <K_*} + \tilde{\epsilon}_{B, k}^{Kn, <K_*} \right) \right)$ as well as all self-interactions of the latter term inside $\overline{\phi_k^{n, <K_*, M, R}}$ are seen to lead to terms vanishing with respect to $\|\cdot\|_{S([0, R])}$ as $n \rightarrow \infty$ on account of their pointwise decay. Moreover, all interactions of $\eta^{Kn, <K_*}$ with u^n are seen to be either vanishing provided we let $n \rightarrow \infty$ using (iv.2), or else of the form $O(\|\eta^{Kn, <K_*}\|_S^2)$.

Step 3 of the proof of Lemma 7.21: *The preceding bound implies the following:*

$$(7.53) \quad \begin{aligned} & \left(\sum_k \left\| (\eta_k^{Kn, <K_*} + \epsilon_{bounded, k}^{<K_*})^{M, R} - \overline{\epsilon_{bounded, k}^{<K_*, M, R}} \right\|_{S_k([0, R])}^2 \right)^{\frac{1}{2}} \\ & \lesssim \delta_4 e^{-R^{1000} - 2K_*} C_{12} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) + R^{41} (\log R + K_*)^{\frac{1}{2}} \|\eta^{Kn, <K_*}\|_{S([0, R])} C_{13} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) \\ & \quad + F \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) \|\eta^{Kn, <K_*}\|_{S([0, R])}^3 + c_R \|\eta^{Kn, <K_*}\|_S. \end{aligned}$$

In fact, we can use the more general observation that if $f_{(k)}$ is a family of functions satisfying

$$\sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_{k \in \mathbb{Z}} \|P_{k+a} f_{(k)}\|_{S_{k+a}([0, R])}^2 \right)^{\frac{1}{2}} \leq A$$

as well as

$$\lim_{n \rightarrow \infty} \sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_{k < \mu_n} \|P_{k+a} f(k)\|_{S_{k+a}([0,R])}^2 \right)^{\frac{1}{2}} = 0,$$

then we have

$$\left(\sum_{k \in \mathbb{Z}} \|\tilde{P}_k[(U_{<k}^{(\phi^n, <K_*)})^\dagger f(k)]\|_{S_k([0,R] \times \mathbb{R}^2)}^2 \right)^{\frac{1}{2}} \lesssim A + c_R \|\eta^{Kn, <K_*}\|_S + o(1)$$

as $n \rightarrow \infty$. To see this, we consider the case $k = 0$ and apply the usual frequency trichotomy to the expression

$$\tilde{P}_0((U_{<0}^{(\phi^n, <K_*)})^\dagger f(0)).$$

Consider for example the low-high case

$$\tilde{P}_0(P_{<-10}(U_{<0}^{(\phi^n, <K_*)})^\dagger P_0 f(0)).$$

Then the estimate for those parts of $\|\cdot\|_{S_0}$ not involving any derivatives is immediate. Next, consider the part of the norm involving the operator $(\partial_t + \partial_r)$. If it falls on $P_0 f(0)$, then the estimate is again immediate. Thus consider now the expression

$$\tilde{P}_0((\partial_t + \partial_r)P_{<-10}(U_{<0}^{(\phi^n, <K_*)})^\dagger P_0 f(0)).$$

Schematically, this reduces to estimating terms of the form

$$\tilde{P}_0((\partial_t + \partial_r)B_{<-10}^{(\phi^n, <K_*)} P_0 f(0)).$$

Call the high-frequency factors in the bilinear expression defining $B_{<-10}^{(\phi^n, <K_*)}$ the “perturbative factors”. Expanding out $B_{<-10}^{(\phi^n, <K_*)}$ using (7.39), we can easily deal with the cases where the perturbative factor is either u^n or a dispersive term

$$\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right).$$

In the former case, we obtain a contribution $o(1)$ due to the diverging frequency supports of $f(0)$ and u^n , and the same applies in the latter case, this time due to the vanishing property of the dispersive terms. To see this latter affirmation, consider a schematic term

$$(7.54) \quad \tilde{P}_0 \left((\partial_t + \partial_r) \sum_{k < -10} \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right) P_0 f(0) \right).$$

Then if there is a low-high frequency interaction inside $\tilde{P}_k(\dots)$ and the derivative $(\partial_t + \partial_r)$ falls on the high-frequency factor, we split the term as follows

$$\begin{aligned} & \tilde{P}_0 \left((\partial_t + \partial_r) \sum_{k < -10} \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right) P_0 f(0) \right) \\ &= \chi_{\{r < \Lambda\}} \tilde{P}_0 \left((\partial_t + \partial_r) \sum_{k < -10} \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right) P_0 f(0) \right) \\ &+ \chi_{\{r \geq \Lambda\}} \tilde{P}_0 \left((\partial_t + \partial_r) \sum_{k < -10} \tilde{P}_k \left((U_{<k}^{(\phi^n, <K_*)})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn, <K_*} + \tilde{\epsilon}_{B,k}^{Kn, <K_*} \right) \right) P_0 f(0) \right). \end{aligned}$$

Then we place the first term on the right into the $(-)$ -component, which is easily seen to be $o(1)$ as $n \rightarrow \infty$ for any $\Lambda > 0$, while we place the second term on the right into the (\pm) -component depending on the type of $(\partial_t + \partial_r)\tilde{\epsilon}_{b,k}^{Kn, <K_*}$ and $(\partial_t + \partial_r)\tilde{\epsilon}_{B,k}^{Kn, <K_*}$. If the latter is in the $(+)$ -case, then we gain smallness by using

$\|\chi_{\{r>\Lambda\}} P_0 f(0)\|_{L_{t,x}^\infty} \lesssim \Lambda^{-\frac{1}{2}} \|f(0)\|_{L_t^\infty \dot{H}_x^1}$. Similarly one deals with the $(-)$ -case. Since we may let $\Lambda \rightarrow \infty$, these contributions are $o(1)$. The remaining frequency interactions inside (7.54) are handled similarly.

Finally, if the perturbative factor in $B_{<-10}^{(\phi^{n,<K_*},R)}$ happens to be $\eta^{Kn,<K_*}$, then the contribution is seen to be bounded by $c_R \|\eta^{Kn,<K_*}\|_S$ with $\lim_{R \rightarrow \infty} c_R = 0$ due to the R -dependence of $B_{<-10}^{(\phi^{n,<K_*},R)}$.

Step 4 of the proof of Lemma 7.21: *The preceding step implies the following:*

$$(7.55) \quad \begin{aligned} & \|\eta^{Kn,<K_*}\|_{S([0,R])} \\ & \lesssim \delta_4 e^{-R^{1000-2K_*}} C_{13} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) + R^{41} (\log R + K_*)^{\frac{1}{2}} \|\eta^{Kn,<K_*}\|_{S([0,R])} C_{14} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) \\ & \quad + E \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) \|\eta^{Kn,<K_*}\|_{S([0,R])}^2 + c_{R,M} \|\eta^{Kn,<K_*}\|_{S([0,R])}, \end{aligned}$$

where

$$\lim_{R,M \rightarrow \infty} c_{R,M} = 0.$$

To get this conclusion, we need to pass from

$$\overline{(\eta_k^{Kn,<K_*} + \epsilon_{bounded,k}^{<K_*})^{M,R}} - \overline{\epsilon_{bounded,k}^{<K_*,M,R}}$$

to $\eta_k^{Kn,<K_*}$. For this it suffices to control

$$\overline{(\eta_k^{Kn,<K_*} + \epsilon_{bounded,k}^{<K_*})^{M,R}} - \overline{\epsilon_{bounded,k}^{<K_*,M,R}} - \eta_k^{Kn,<K_*},$$

which consists of cubic terms two of which (the high frequency and differentiated low frequency term) are “perturbative”. But then, exploiting the precise definition of this expression, we easily see that its square-sum norm may be bounded by $c_{R,M} \|\eta^{Kn,<K_*}\|_S$.

Finally, we have reached the point where we can get the conclusion of the lemma via a bootstrap argument:

Step 5 of the proof of Lemma 7.21: *Conclusion of the proof via time localization.* Observe that we can choose R larger than any function of $\|\epsilon_{bounded}^{<K_*}\|_S$ and $\|u^n\|_S$. In particular, all terms on the right hand side of (7.55) except the second one are perturbative. However, this term first arose in Step 1 of the proof of Lemma 7.21 on account of the norms of certain source terms in $F^{(R,M)}$ with good divisibility properties, and propagated through the following steps. In particular, we can replace the bad factor $R^{41} (\log R + K_*)^{\frac{1}{2}} C_{10} (\|\epsilon_{bounded}^{<K_*}\|_S)$ by a good one, such as R^{-1} , by subdividing the time interval $[0, R]$ into

$$M_1 := C_{10} (\|\epsilon_{bounded}^{<K_*}\|_S) R^{100} (\log R + K_*)^{100}$$

many sub-intervals I_1, I_2, \dots, I_{M_1} , say, and replacing $[0, R] \times \mathbb{R}^2$ by $I_j \times \mathbb{R}^2$. Doing so then allows us to infer the better estimate

$$(7.56) \quad \begin{aligned} \|\eta^{Kn,<K_*}\|_{S[I_1]} & \lesssim \delta_4 e^{-R^{1000-2K_*}} C_{15} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) + R^{-1} \|\eta^{Kn,<K_*}\|_{S[I_1]} \\ & \quad + F \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) \|\eta^{Kn,<K_*}\|_{S[I_1]}^2 + c_{R,M} \|\eta^{Kn,<K_*}\|_{S[I_1]}, \end{aligned}$$

from which we infer

$$\|\eta^{Kn,<K_*}\|_{S[I_1]} \lesssim \delta_4 e^{-R^{1000-2K_*}} C_{16} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right).$$

Progressing inductively through the intervals I_j , $j = 2, 3, \dots, M_1$, and each time using the improved bound for the previous interval for the initial data, we finally infer the bound

$$\|\eta^{Kn,<K_*}\|_{S([0,R])} \lesssim \delta_4 e^{C_{17} (\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S)} R^{100} (\log R + K_*)^{100} e^{-R^{1000-2K_*}} C_{18} \left(\|\epsilon_{bounded}^{<K_*}\|_S, \|u^n\|_S \right) \lesssim \delta_4 \ll \delta_3,$$

if we choose R sufficiently large and δ_4 sufficiently small compared to δ_3 . This then proves Lemma 7.21 up to establishing *Observation 1* and *Observation 2*, which we now do. \square

Proof of Observation 1. Write $\epsilon \equiv \epsilon_{\text{bounded}}^{<K_*}$ and $\phi \equiv \phi^{n, <K_*}$. We consider the quantity $U_{<s}^{(\phi, R)}(U_{<s}^{(\epsilon, R)})^\dagger$. Differentiating it with respect to s , we find

$$\partial_s \left(U_{<s}^{(\phi, R)}(U_{<s}^{(\epsilon, R)})^\dagger \right) = U_{<s}^{(\phi, R)}(B_s^{(\phi, R)} - B_s^{(\epsilon, R)})(U_{<s}^{(\epsilon, R)})^\dagger$$

and so since $\lim_{s \rightarrow -\infty} U_{<s}^{(\phi, R)}(U_{<s}^{(\epsilon, R)})^\dagger = Id$, we get

$$U_{<s}^{(\phi, R)}(U_{<s}^{(\epsilon, R)})^\dagger = \int_{-\infty}^s U_{<\tilde{s}}^{(\phi)}(B_{\tilde{s}}^{(\phi, R)} - B_{\tilde{s}}^{(\epsilon, R)})(U_{<\tilde{s}}^{(\epsilon, R)})^\dagger d\tilde{s} + Id.$$

We shall then simply set

$$G_k^{(n)} := U_{<k}^{(\phi, R)}(U_{<k}^{(\epsilon, R)})^\dagger(0, 0).$$

To see that this works, note that by our definition of $U_{<k}^{(\phi, R)}$, we get

$$\begin{aligned} \left(U_{<k}^{(\phi, R)}(U_{<k}^{(\epsilon, R)})^\dagger - G_k^{(n)} \right) U_{<k}^{(\epsilon, R)} \epsilon_k &= \left(\int_0^R \int_{-\infty}^k \partial_t \left(U_{<\tilde{s}}^{(\phi)}(B_{\tilde{s}}^{(\phi, R)} - B_{\tilde{s}}^{(\epsilon, R)})(U_{<\tilde{s}}^{(\epsilon, R)})^\dagger \right) d\tilde{s} dt \right) U_{<k}^{(\epsilon, R)} \epsilon_k \\ &\quad + \left(U_{<k}^{(\phi, R)}(0, \cdot)(U_{<k}^{(\epsilon, R)})^\dagger(0, \cdot) - G_k^{(n)} \right) U_{<k}^{(\epsilon, R)} \epsilon_k. \end{aligned}$$

Then by expanding the difference $(B_{\tilde{s}}^{(\phi, R)} - B_{\tilde{s}}^{(\epsilon, R)})$ one checks that for the first term on the right we have

$$\sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_{k \in \mathbb{Z}} \|P_{k+a}(\cdot)\|_{S_{k+a}}^2 \right)^{\frac{1}{2}} \leq o(1) + c_R \|\eta^{Kn, <K_*}\|_S$$

as $n \rightarrow \infty$, where $\lim_{R \rightarrow \infty} c_R = 0$. Furthermore, split

$$\epsilon_k = \chi_{B_{R_*}} \epsilon_k + (1 - \chi_{B_{R_*}}) \epsilon_k$$

for some $R_* \gg R$ as in case (v) of Step 1 of the proof of Lemma 7.21 and write

$$\begin{aligned} &\left(U_{<k}^{(\phi, R)}(0, r)(U_{<k}^{(\epsilon, R)})^\dagger(0, r) - G_k^{(n)} \right) U_{<k}^{(\epsilon, R)} \chi_{B_{R_*}} \epsilon_k \\ &= \left(\int_0^r \int_{-\infty}^k \partial_{\tilde{r}} \left(U_{<\tilde{s}}^{(\phi)}(B_{\tilde{s}}^{(\phi, R)} - B_{\tilde{s}}^{(\epsilon, R)})(U_{<\tilde{s}}^{(\epsilon, R)})^\dagger \right)(0, \tilde{r}) d\tilde{s} d\tilde{r} \right) U_{<k}^{(\epsilon, R)} \chi_{B_{R_*}} \epsilon_k. \end{aligned}$$

Then arguing as in (v) of Step 1 of the proof of Lemma 7.21, one gets that the preceding expression satisfies

$$\sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_{k \in \mathbb{Z}} \|P_{k+a}(\cdot)\|_{S_{k+a}}^2 \right)^{\frac{1}{2}} \leq o(1) + \delta_4 e^{-R^{1000} - 2K_*}$$

for K, n large enough, and the same conclusion applies to the contribution of $(1 - \chi_{B_{R_*}}) \epsilon_k$ for R_* large enough relative to R, K_* . This proves the first part of *Observation 1*, and the second part is obtained similarly, using divisibility for $F^{(R, M)}(\epsilon)$. \square

Proof of Observation 2. We write

$$U_{<s}^{(\phi, R)} = \int_{-\infty}^s U_{<\tilde{s}}^{(\phi, R)} B_{\tilde{s}}^{(\phi, R)} d\tilde{s}, \quad U_{<s}^{(u^n, R)} = \int_{-\infty}^s U_{<\tilde{s}}^{(u^n, R)} B_{\tilde{s}}^{(u^n, R)} d\tilde{s},$$

whence we get (specializing to frequency $k = 0$)

$$\begin{aligned} (U_{<0}^{(\phi, R)} - U_{<0}^{(u^n, R)}) \overline{u_0^{n, M, R}} &= \int_{-\infty}^0 U_{<\tilde{s}}^{(\phi, R)} (B_{\tilde{s}}^{(\phi, R)} - B_{\tilde{s}}^{(u^n, R)}) \overline{u_0^{n, M, R}} d\tilde{s} \\ &\quad + \int_{-\infty}^0 (U_{<\tilde{s}}^{(\phi, R)} - U_{<\tilde{s}}^{(u^n, R)}) B_{\tilde{s}}^{(u^n, R)} \overline{u_0^{n, M, R}} d\tilde{s}. \end{aligned}$$

One then reiterates in the second expression on the right, and so it suffices to bound the first term on the right. Then recalling that μ_n denotes the frequency dividing between the lowest frequency non-atomic constituent and the first large frequency atom, then if $0 < \mu_n$ we can close by exploiting the exponential decay of all

of the $\epsilon_{bounded,s_1}^K$ etc. as well as the bound in (iv.2) to control $\eta_{s_1}^{nK}$ for $s_1 < \mu_n$. On the other hand, if $0 \geq \mu_n$, we split

$$\begin{aligned} \int_{-\infty}^0 U_{<\tilde{s}}^{(\phi,R)} (B_{\tilde{s}}^{(\phi,R)} - B_{\tilde{s}}^{(u^n,R)}) \overline{u_0^{n,M,R}} d\tilde{s} &= \int_{-\infty}^{\mu_n} U_{<\tilde{s}}^{(\phi,R)} (B_{\tilde{s}}^{(\phi,R)} - B_{\tilde{s}}^{(u^n,R)}) \overline{u_0^{n,M,R}} d\tilde{s} \\ &\quad + \int_{\mu_n}^0 U_{<\tilde{s}}^{(\phi,R)} (B_{\tilde{s}}^{(\phi,R)} - B_{\tilde{s}}^{(u^n,R)}) \overline{u_0^{n,M,R}} d\tilde{s}. \end{aligned}$$

The first term on the right is bounded by exploiting the exponential decay of each term in

$$\epsilon_{bounded,s_1}^{<K_*} + \sum_{b=1}^B \tilde{P}_{s_1} \left((U_{<s_1}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{b,s_1}^{Kn,<K_*} + (U_{<s_1}^{(\phi^{Kn,<K_*})})^\dagger \tilde{\epsilon}_{B,s_1}^{Kn,<K_*} \right) + \eta_{s_1}^{Kn,<K_*}$$

with respect to $s_1 - \mu_n$, while in the second term one loses μ_n due to the integration over \tilde{s} , but gains $o(1) 2^{-\sigma\mu_n}$ due to the smallness of $\overline{u_0^n}$. \square

Continuing with the Stage 1 of the proof of Theorem 7.17, we pass to

Step 2: We adjust the decomposition (7.39) slightly and thereby achieve control on the interval $[R, t_1^n - T_*]$. Specifically, we replace (7.39) by

$$(7.57) \quad \phi^{n,<K_*} = u^n + \sum_k \tilde{P}_k \left((U_{<k}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_k^{Kn,<K_*} \right) + \sum_k \tilde{P}_k \left((U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right) + \tilde{\eta}^{Kn,<K_*},$$

where we let $\tilde{\epsilon}_k^{Kn,<K_*}$ be the evolution with respect to (7.34) of the initial data

$$\left((U_{<k}^{(\phi^{n,<K_*})}) P_{[-K,\infty)} \epsilon_{bounded,k}^{<K_*} \partial_t (U_{<k}^{(\phi^{n,<K_*})}) P_{[-K,\infty)} \epsilon_{bounded,k}^{<K_*} + (U_{<k}^{(\phi^{n,<K_*})}) \partial_t P_{[-K,\infty)} \epsilon_{bounded,k}^{<K_*} \right) \Big|_{t=R}.$$

Observe that we have included an extra low frequency cutoff $P_{[-K,\infty)}$ for the data and we absorb the corresponding correction into $\tilde{\eta}^{Kn,<K_*}$. Then it is straightforward to check that

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \left\| \nabla_{t,x} (\tilde{\eta}^{Kn,<K_*} - \eta^{Kn,<K_*})(R, \cdot) \right\|_{L_x^2} = 0.$$

Now we can formulate the following important

Lemma 7.22. *There is T_* sufficiently large (depending on the first temporally unbounded profile) such that if K, R, n are sufficiently large and $B = B(\|u^n\|_S, \delta_2)$ is sufficiently large, then we have*

$$\left\| \tilde{\eta}^{Kn,<K_*} \right\|_{S([R, t_1^n - T_*])} < \delta_2.$$

Furthermore, there exist (not necessarily frequency localized) functions $H_{(k)}^{(1)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*})$, $H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*})$ and $g_{(k)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*})$, $k \in \mathbb{Z}$, which are multilinear expressions in its frequency localized inputs as well as in the gauge transformations $U_{<h}^{(\phi^{n,<K_*})}$ and $(U_{<h}^{(\phi^{n,<K_*})})^\dagger$, and with

$$\sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_k \left\| P_{k+a} H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}, u^n) \right\|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}} \leq c(\eta_{A^n}(\tilde{\epsilon}_B^{Kn,<K_*}))$$

with $c(\gamma) \rightarrow 0$ as $\gamma \rightarrow 0$, as well as with

$$\begin{aligned} \sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_k \left\| P_{k+a} H_{(k)}^{(1)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right\|_{L_t^1 L_x^2}^2 \right)^{\frac{1}{2}} &\lesssim \left(\left\| \tilde{\eta}^{Kn,<K_*} \right\|_S + \left\| \tilde{\eta}^{Kn,<K_*} \right\|_S^{100} \right) F(\left\| \tilde{\epsilon}_B^{Kn,<K_*} \right\|_S), \\ \sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_k \left\| P_{k+a} g_{(k)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right\|_{S_{k+a}}^2 \right)^{\frac{1}{2}} &\lesssim \left\| \tilde{\epsilon}_B^{Kn,<K_*} \right\|_S^2 + \left\| \tilde{\eta}^{Kn,<K_*} \right\|_S^2, \end{aligned}$$

and correction terms κ_k , $k \in \mathbb{Z}$ with

$$\sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_k \|P_{k+a} \kappa_k\|_S^2 \right)^{\frac{1}{2}} \leq o(1)$$

as $R, T_*, n \rightarrow \infty$, and such that

$$(7.58) \quad \tilde{\square}_{A^n}(U_{<k-10}^{(\phi^{n,<K_*})} \tilde{\eta}_k^{Kn,<K_*} + g_{(k)} + \kappa_k) = H_{(k)}^{(1)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) + H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) + o^{(k)}(1),$$

where the error satisfies the relation

$$\lim_{R, T_*, n \rightarrow \infty} \sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_k \|P_{k+a} o^{(k)}(1)\|_{L_t^1 L_x^2([R, t_1^n - T_*])}^2 \right)^{\frac{1}{2}} = 0.$$

Proof. Consider

$$(7.59) \quad \tilde{\square}_{A^n} \left(U_{<k-10}^{(\phi^{n,<K_*})} \overline{\phi_k^{n,<K_*}} - U_{<k-10}^{(u^n)} \overline{u_k^n} - \tilde{\epsilon}_k^{Kn,<K_*} - \left(\sum_{b=1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right).$$

Then on the one hand this equals

$$\tilde{\square}_{A^n} \left(U_{<k-10}^{(\phi^{n,<K_*})} \tilde{\eta}_k^{Kn,<K_*} + \zeta_k \right),$$

where (setting $\tilde{\epsilon}_{0,k}^{Kn,<K_*} := \tilde{\epsilon}_k^{Kn,<K_*}$ for simplicity)

$$\begin{aligned} \zeta_k &= (U_{<k-10}^{(\phi^{n,<K_*})} - U_{<k-10}^{(u^n)}) \overline{u_k^n} \\ &\quad + U_{<k-10}^{(\phi^{n,<K_*})} \tilde{P}_k \left((U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=0}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right) - \left(\sum_{b=0}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \\ &\quad + U_{<k-10}^{(\phi^{n,<K_*})} \left(\overline{\phi_k^{n,<K_*}} - \overline{u_k^n} - \tilde{P}_k \left((U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=0}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right) - \tilde{\eta}_k^{Kn,<K_*} \right), \end{aligned}$$

is of the form $\kappa_k + g_{(k)}$ as claimed. In fact, for the first difference term on the right, it is of type κ_k due to a small variation on *Observation 2* in the preceding step, while for the second difference term on the right, i.e.

$$(7.60) \quad U_{<k-10}^{(\phi^{n,<K_*})} \tilde{P}_k \left((U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=0}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right) - \left(\sum_{b=0}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right),$$

one observes that it coincides with

$$(7.61) \quad \left(\sum_{b=0}^B \tilde{P}_k \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{P}_k \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) - \left(\sum_{b=0}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right),$$

up to the sum of a term satisfying the desired bound for κ_k and a term of type $g_{(k)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*})$. To see this, note that one can write the term (7.60) schematically as

$$\nabla_x (U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=0}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right)$$

and expanding out $\nabla_x (U_{<k-10}^{(\phi^{n,<K_*})})^\dagger$ using the equation for $U_{<k-10}^{(\phi^{n,<K_*})}$, one retains only those terms involving $\tilde{\eta}^{Kn,<K_*}$ and $\tilde{\epsilon}_B^{Kn,<K_*}$ in combination with the factor $\tilde{\epsilon}_{B,k}^{Kn,<K_*}$ to be included into the term $g_{(k)}$.

The preceding difference (7.61) also satisfies the desired bound for κ_k as one easily sees after applying the operator \tilde{P}_k to the equation (7.34). Finally, for the third term in the above formula for ζ_k we note that it arises due to the nonlinear interactions inside $\overline{\phi_k^{n,<K_*}}$, and its contribution to the norm controlling κ_k is easily seen to be bounded by $O(\|\tilde{\eta}^{Kn,<K_*}\|_S^2 + \|\tilde{\epsilon}_B^{Kn,<K_*}\|_S^2)$ in the limit $R, T, n \rightarrow \infty$.

Getting back to (7.59), on the other hand we can write this as

$$\begin{aligned}
(7.62) \quad & \tilde{\square}_{A^n} \left(U_{<k-10}^{(\phi^{n,<K_*})} \overline{\phi_k^{n,<K_*}} - U_{<k-10}^{(u^n)} \overline{u_k^n} - \left(\sum_{b=0}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right) \\
&= \tilde{\square}_{A^n} \left(U_{<k-10}^{(\phi^{n,<K_*})} \overline{\phi_k^{n,<K_*}} - U_{<k-10}^{(u^n)} \overline{u_k^n} \right) \\
&= \square \left(U_{<k-10}^{(\phi^{n,<K_*})} \overline{\phi_k^{n,<K_*}} \right) - \square \left(U_{<k-10}^{(u^n)} \overline{u_k^n} \right) + U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \left(U_{<k-10}^{(u^n)} \overline{u_k^n} - U_{<k-10}^{(\phi^{n,<K_*})} \overline{\phi_k^{n,<K_*}} \right).
\end{aligned}$$

As for the last term on the right, expand it out as

$$\begin{aligned}
& U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \left(U_{<k-10}^{(u^n)} \overline{u_k^n} - U_{<k-10}^{(\phi^{n,<K_*})} \overline{\phi_k^{n,<K_*}} \right) \\
&= U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \left((U_{<k-10}^{(u^n)} - U_{<k-10}^{(\phi^{n,<K_*})}) \overline{u_k^n} \right) \\
&\quad - U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \left(U_{<k-10}^{(\phi^{n,<K_*})} (\overline{\phi_k^{n,<K_*}} - \overline{u_k^n}) \right).
\end{aligned}$$

Here the first term on the right is easily seen to be of the form

$$\begin{aligned}
(7.63) \quad & U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \left((U_{<k-10}^{(u^n)} - U_{<k-10}^{(\phi^{n,<K_*})}) \overline{u_k^n} \right) \\
&= U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \left((U_{<k-10}^{(u^n)} - U_{<k-10}^{(u^n+\tilde{\eta}^{Kn,<K_*})}) \overline{u_k^n} \right) + o^{(k)}(1),
\end{aligned}$$

where the error $o^{(k)}(1)$ has the claimed vanishing property provided $T_*, R, n \rightarrow \infty$, and similarly we have

$$\begin{aligned}
(7.64) \quad & U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \left(U_{<k-10}^{(\phi^{n,<K_*})} (\overline{\phi_k^{n,<K_*}} - \overline{u_k^n}) \right) \\
&= U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \left(U_{<k-10}^{(\phi^{n,<K_*})} \tilde{\eta}_k^{Kn,<K_*} \right) \\
&\quad + U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha \left(U_{<k-10}^{(\phi^{n,<K_*})} P_k \left(\sum_{b=0}^B (U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{b,k}^{Kn,<K_*} + (U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right) \\
&\quad + o^{(k)}(1).
\end{aligned}$$

Of the preceding two expressions (7.63) and (7.64), the most delicate case appears when the derivative ∂^α in (7.63) falls on the term u_k^n , since then smallness has to come from essentially a factor of the form $\tilde{\eta}_{<k-10}^{Kn,<K_*}$ arising essentially from $U_{<k-10}^{(u^n)} - U_{<k-10}^{(u^n+\tilde{\eta}^{Kn,<K_*})}$. Dealing with this situation requires an energy-type bound for $\tilde{\eta}_k^{Kn,<K_*}$ for frequencies $k \leq \mu_n$, analogous to Lemma 7.10. In fact, making a suitable bootstrap hypothesis for the very low frequency part of $\tilde{\eta}_k^{Kn,<K_*}$ and proceeding as in the proof of Lemma 7.10, we infer the following bound

$$\left\| \nabla_{t,x} \tilde{\eta}_k^{Kn,<K_*} \right\|_{L_t^\infty L_x^2} \leq o(1) \cdot 2^{-\sigma|k-\mu_n|},$$

where $o(1)$ vanishes as $n \rightarrow \infty$. Using this and, as usual, splitting into the cases $k \leq \mu_n$ and $k > \mu_n$, one easily infers that as $n \rightarrow \infty$,

$$\sum_{a \in \mathbb{Z}} 2^{\sigma|a|} \left(\sum_k \left\| P_{k+a} \left(U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha ((U_{<k-10}^{(u^n)} - U_{<k-10}^{(u^n+\tilde{\eta}^{Kn,<K_*})}) u_k^n) \right) \right\|_{L_t^1 L_x^2([R, t_1^* - T_*])} \right)^{\frac{1}{2}} = o(1)$$

On the other hand, the first two lines on the right in (7.64) cancel against a corresponding term obtained when expanding

$$\square(U_{<k-10}^{(\phi^{n,<K_*})} \overline{\phi_k^{n,<K_*}}) - \square(U_{<k-10}^{(u^n)} \overline{u_k^n}) = U_{<k-10}^{(\phi^{n,<K_*})} F_k(\phi^{n,<K_*} \nabla_{t,x} \phi^{n,<K_*}) - U_{<k-10}^{(u^n)} F_k(u^n, \nabla_{t,x} u^n).$$

In fact, we get

$$\begin{aligned}
& U_{<k-10}^{(\phi^{n,<K_*})} F_k(\phi^{n,<K_*}, \nabla_{t,x} \phi^{n,<K_*}) - U_{<k-10}^{(u^n)} F_k(u^n, \nabla_{t,x} u^n) \\
&= U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha (U_{<k-10}^{(\phi^{n,<K_*})} \tilde{\eta}_k^{Kn,<K_*}) \\
(7.65) \quad &+ U_{<\mu_n}^{(u^n)} \mathcal{A}_{\alpha,low}(u^n) (U_{<\mu_n}^{(u^n)})^\dagger \partial^\alpha P_k \left(\sum_{b=0}^B (U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{b,k}^{Kn,<K_*} + (U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \\
&+ H_{(k)}^{(1)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{Kn,<K_*}) + H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{Kn,<K_*}) + o^{(k)}(1),
\end{aligned}$$

where the error $o^{(k)}(1)$ is again as in the statement of Lemma 7.22. In fact, with the exception of the second term on the right in (7.65), all terms on the left in (7.65) involving at least one perturbative factor of the form

$$\sum_{b=0}^B (U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{b,k}^{Kn,<K_*}$$

are seen to be of the form $o^{(k)}(1)$ by using Lemma 7.20. The same applies for all terms involving at least one perturbative factor u^n . For the remaining interactions, the terms of the form $H_{(k)}^{(1)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{Kn,<K_*})$ arise when at least one of the perturbative terms in the source is of the form $\tilde{\eta}^{Kn,<K_*}$ while all other perturbative factors are of the form

$$\sum_k \tilde{P}_k((U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{B,k}^{Kn,<K_*}),$$

while the terms of the form $H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{Kn,<K_*})$ arise when all perturbative factors are of the form

$$\sum_k (U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{Bk}^{Kn,<K_*}.$$

Now a priori in expressions of the latter kind the norm of the expression will also depend on $\|\phi^{Kn,<K_*}\|_S$, provided a derivative falls on the factor $U_{<k-10}^{(\phi^{n,<K_*})}$. In this case we re-express $\phi^{Kn,<K_*}$ using (7.57). On the other hand, all terms in $H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{Kn,<K_*})$ where no derivative lands on $U_{<k-10}^{(\phi^{n,<K_*})}$ are left unchanged. This process can be continued, absorbing terms either into $H_{(k)}^{(1)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{Kn,<K_*})$ or into $H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{Kn,<K_*})$ or into $o^{(k)}(1)$, until all remaining terms in $H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{Kn,<K_*})$ in which a derivative can still potentially fall on a factor $(U_{<k-10}^{(\phi^{n,<K_*})})^\dagger$ can be estimated purely in terms of Strichartz norms of its inputs. Then it is easy to see that it suffices to control $\|\nabla_{t,x}(U_{<k-10}^{(\phi^{n,<K_*})})^\dagger\|_{L_t^\infty L_x^{2+}}$, which we do by energy conservation. This proves the relation (7.58) and a straightforward bootstrap/divisibility argument then yields the bound

$$\|\tilde{\eta}^{Kn,<K_*}\|_{S([R,t_1^n-T_*])} < \delta_2,$$

provided we have picked R, T_* , and n sufficiently large. In fact, to pass from the bound on $\{U_{<k-10}^{(\phi^{n,<K_*})} \tilde{\eta}_k^{Kn,<K_*}\}_{k \in \mathbb{Z}}$ to the bound for $\tilde{\eta}^{Kn,<K_*}$, one argues exactly as in Step 3 of the proof of the preceding Lemma 7.21. \square

At this point, we have shown that under the hypotheses of Theorem 7.17, we control the solution $\phi^{n,<K_*}$ for n sufficiently large on the full interval $[0, t_1^n - T_*]$ for a fixed but sufficiently large T_* , and that we in fact obtain bounds on $\|\phi^{n,<K_*}\|_S$ which are independent of K_* and n .

Our aim now is to continue the preceding process on the next large time interval $[t_1^n - T_*, t_2^n - T_{1*}]$ for T_{1*} sufficiently large.

Stage 2: *Control of the solution on $[t_1^n - T_*, t_2^n - T_{1*}]$.* Here we want to repeat the method of Stage 1 by replacing the initial time $t = 0$ by the initial time $t = t_1^n - T_*$. A basic technical difficulty we face here is that we first have to introduce the analogue of $\epsilon_{bounded}^{<K_*}$, say $\epsilon_{1bounded}^{<K_*}$, by suitably modifying the function

$\sum_k (U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{1,k}^{Kn,<K_*}$, i.e. the profile with $b = 1$, which concentrates at time $t = t_1^n - T_*$. For this we will also have to take into account the perturbation $\tilde{\eta}^{Kn,<K_*}$, which, while small in energy, may lead to a significant effect on the L_x^∞ -norm. To construct the profile, we more or less proceed as in the proof of Lemma 7.14, but the following lemma simplifies things.

Lemma 7.23. *The function $\nabla_{t,x} \tilde{\eta}^{Kn,<K_*}(t_1^n - T_*, \cdot)$ vanishes weakly in $L_x^2(\mathbb{R}^2)$ as $T_*, K, n \rightarrow \infty$.*

Proof. This is a consequence of the identity (7.58) in the preceding Lemma 7.22. Writing the flow of (7.34) associated with the data $(0, f)$ at $t = 0$ in the form

$$U_{A^n}(t)(f),$$

we have

$$\begin{aligned} & \left(U_{<k-10}^{(\phi^{n,<K_*})} \tilde{\eta}_k^{Kn,<K_*} + g_{(k)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) \Big|_{t_1^n - T_*} \\ &= S_{A^n} \left(\left(U_{<k-10}^{(\phi^{n,<K_*})} \tilde{\eta}_k^{Kn,<K_*} + g_{(k)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) [0] \right) \Big|_{t_1^n - T_*} \\ &+ \int_0^t U_{A^n}(t-s) \left(H_{(k)}^{(1)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) + H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) (s, \cdot) ds \Big|_{t_1^n - T_*} + o(1), \end{aligned}$$

where the error vanishes (in the energy sense) upon letting $T_*, K, n \rightarrow \infty$. Consider the principal contribution to the last integral, which is the expression

$$\int_0^t U_{A^n}(t-s) \left(H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) (s, \cdot) ds \Big|_{t_1^n - T_*}.$$

Recall that $H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*})$ is a multilinear expression with $\sum_k (U_{<k-10}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{B,k}^{Kn,<K_*}$ in its “perturbative factors”. In turn, $\tilde{\epsilon}_{B,k}^{Kn,<K_*}$ can be expanded as a sum of free waves associated to (7.34) up to an arbitrarily small error $\tilde{\epsilon}_{B_1}^{Kn,<K_*}$ and each of these free waves concentrates at a time scale diverging from $t_1^n - T_*$. This means that up to an arbitrarily small error and upon letting $n \rightarrow \infty$, the function

$$\int_0^t U_{A^n}(t-s) \left(H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) (s, \cdot) ds \Big|_{t_1^n - T_*}$$

converges in the S -norm to a solution of (7.34) concentrating away from $t = t_1^n - T_*$, and in particular we have

$$\nabla_{t,x} \int_0^t U_{A^n}(t-s) \left(H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) (s, \cdot) ds \Big|_{t_1^n - T_*} \rightarrow 0.$$

As for the remaining integral

$$\int_0^t U_{A^n}(t-s) \left(H_{(k)}^{(1)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) (s, \cdot) ds \Big|_{t_1^n - T_*},$$

inserting the principal term for $\tilde{\eta}^{Kn,<K_*}$, which equals

$$\int_0^t U_{A^n}(t-s) \left(H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) (s, \cdot) ds,$$

results in a similar term, again vanishing weakly at $t = t_1^n - T_*$ by the same kind of argument, and one winds up with another error term stemming from substituting

$$\int_0^t U_{A^n}(t-s) \left(H_{(k)}^{(1)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) (s, \cdot) ds$$

for $\tilde{\eta}^{Kn,<K_*}$. In light of the smallness of $\|\tilde{\epsilon}_B^{Kn,<K_*}\|_S$ and $\|\tilde{\eta}^{Kn,<K_*}\|_S$, this term has much smaller S -norm than

$$\int_0^t U_{A^n}(t-s) \left(H_{(k)}^{(1)}(\tilde{\eta}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) (s, \cdot) ds.$$

Repeating this argument a finite number of times leads to a sequence of terms vanishing weakly at $t = t_1^n - T_*$ up to an arbitrarily small error, which yields the claim. \square

We now construct a good bounded profile $\epsilon_{1\text{bounded}}^{<K_*}$ at time $t = t_1^n - T_*$. Departing from (7.57), we get upon restriction to a suitable subsequence with respect to n ,

$$\left(\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{1,k}^{Kn,<K_*} \right) + \tilde{\eta}^{Kn,<K_*} + p_{1*} \right) \Big|_{t_1^n - T_*} = \phi^{n,<K_*} \Big|_{t_1^n - T_*} + o_{L_{loc}^\infty}(1)$$

with the error vanishing as $n \rightarrow \infty$ in the sense of L_{loc}^∞ . Here $p_{1*} \in \mathbb{S}^m$ may be set equal to $\lim_{n \rightarrow \infty} u^n(t_1^n - T_*, 0)$, the limit existing upon passing to a suitable subsequence. In fact, the remaining terms

$$\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=0, b \neq 1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right) \Big|_{t_1^n - T_*}$$

all converge to zero in L_{loc}^∞ as $n \rightarrow \infty$ due to the fact that we have essentially (up to exponentially decaying tails) localized their frequency supports to compact intervals, and each of them converges weakly to zero at time $t_1^n - T_*$ as $n \rightarrow \infty$.

Now due to the essentially compact frequency support of the expression, upon letting $K, n \rightarrow \infty$ along a suitable subsequence, we can find $\tilde{\epsilon}_{1\text{bounded}}^{<K_*}$ such that

$$\left(\sum_k \tilde{P}_k \left((U_{<k}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{1,k}^{Kn,<K_*} \right) + \tilde{\eta}^{Kn,<K_*} \right) \Big|_{t_1^n - T_*} \rightharpoonup \tilde{\epsilon}_{1\text{bounded}}^{<K_*} \Big|_{t_1^n - T_*}$$

and also strongly in H_{loc}^{1+} , in particular in L_{loc}^∞ . In particular, it follows that $\tilde{\epsilon}_{1\text{bounded}}^{<K_*} + p_{1*}$ maps into \mathbb{S}^m . Thus it is natural to set

$$\epsilon_{1\text{bounded}}^{<K_*} \Big|_{t_1^n - T_*} := \tilde{\epsilon}_{1\text{bounded}}^{<K_*} \Big|_{t_1^n - T_*} + p_{1*},$$

and to pass from (7.57) to a new decomposition with a new error term $\tilde{\eta}^{Kn,<K_*}$, where now $\epsilon_{1\text{bounded}}^{<K_*}$ denotes the nonlinear evolution associated with the data $\epsilon_{1\text{bounded}}^{<K_*} \Big|_{t_1^n - T_*}$,

$$(7.66) \quad \phi^{n,<K_*} = u^n - p_{1*} + \epsilon_{1\text{bounded}}^{<K_*} + \sum_k \tilde{P}_k \left((U_{<k}^{(\phi^{n,<K_*})})^\dagger \left(\sum_{b=0, b \neq 1}^B \tilde{\epsilon}_{b,k}^{Kn,<K_*} + \tilde{\epsilon}_{B,k}^{Kn,<K_*} \right) \right) + \tilde{\eta}^{Kn,<K_*}.$$

This is the analogue of (7.38) which we expect to be valid on a large but finite time interval $[t_1^n - T_*, t_1^n - T_* + R_1]$, say, where R_1 will now depend on the scattering properties of $\epsilon_{1\text{bounded}}^{<K_*}$, just as R before was chosen depending on the scattering properties of the evolution of $\epsilon_{\text{bounded}}^{<K_*}$. Also, comparing (7.66) and (7.57), we have

$$\tilde{\eta}^{Kn,<K_*} \Big|_{t_1^n - T_*} = \sum_k (U_{<k}^{(\phi^{n,<K_*})})^\dagger \tilde{\epsilon}_{1,k}^{Kn,<K_*} \Big|_{t_1^n - T_*} + \tilde{\eta}^{Kn,<K_*} \Big|_{t_1^n - T_*} - \tilde{\epsilon}_{1\text{bounded}}^{<K_*}.$$

Then by definition of $\tilde{\epsilon}_{1\text{bounded}}^{<K_*}$ we have $\lim_{n \rightarrow \infty} \left\| \tilde{\eta}^{Kn,<K_*} \Big|_{t_1^n - T_*} \right\|_{L_{loc}^\infty} = 0$, which is analogous to the property of $\eta^{Kn,<K_*}$ in Lemma 7.14. Also, note that for any $R_* > 0$ we have

$$\lim_{n \rightarrow \infty} \left\| \nabla_{t,x} \tilde{\eta}^{Kn,<K_*} \Big|_{t_1^n - T_*} \right\|_{L_x^2(B_{R_*})} = 0.$$

We further observe that due to Lemma 7.23 and its proof we have the following. Denoting by $\tilde{\eta}_{(1)k}^{Kn,<K_*}$ the solution of the approximation to (7.58) given by

$$(7.67) \quad \begin{aligned} & \tilde{\square}_{A^n} \left(U_{<k-10}^{(\phi^{n,<K_*})} \tilde{\eta}_{(1)k}^{Kn,<K_*} + g_{(k)}(\tilde{\eta}_{(1)}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \right) \\ &= H_{(k)}^{(1)}(\tilde{\eta}_{(1)}^{Kn,<K_*}, \tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) + H_{(k)}^{(2)}(\tilde{\epsilon}_B^{Kn,<K_*}, \phi^{n,<K_*}) \end{aligned}$$

with the same initial data as $\tilde{\eta}^{Kn, < K_*}$ at $t = R$, then $\tilde{\eta}_{(1)}^{Kn, < K_*}$ obeys the same bound as $\tilde{\eta}^{Kn, < K_*}$, and moreover we have

$$\lim_{R, T_*, K, n \rightarrow \infty} \left\| \nabla_{t,x} (\tilde{\eta}_{(1)}^{Kn, < K_*} \Big|_{t_1^n - T_*} - \tilde{\eta}^{Kn, < K_*} \Big|_{t_1^n - T_*}) \right\|_{L_x^2} = 0.$$

We also note that

$$\lim_{R, T_*, K, n \rightarrow \infty} \left\| \tilde{\eta}_{(1)}^{Kn, < K_*} - \tilde{\eta}^{Kn, < K_*} \right\|_{S([R, t_1^n - T_*])} = 0.$$

At this point, we briefly pause to make an important

Remark 7.24. *Note that at this stage of the proof of Theorem 7.17 we have already introduced three different perturbation terms $\eta^{Kn, < K_*}$, $\tilde{\eta}^{Kn, < K_*}$, and $\tilde{\eta}_{(1)}^{Kn, < K_*}$. As we intend to iterate the preceding procedure on each of the time intervals $[t_b^n - T_{b*}, t_{b+1}^n - T_{b+1*}]$, $b = 1, 2, \dots, B$, we have to be careful that this process does not result in accruing more and more errors, i.e. we have to ensure uniform smallness of these perturbations. The idea here is that we have the parameters R, T_*, K, n and analogous ones for the later stages (B in total) at our disposal, while B will be chosen as already mentioned only depending on an absolute small constant δ_2 as well as on $\|u^n\|_S$. Thus, by eventually picking R, T_*, K, n etc. sufficiently large, we will be able to ensure that passing from one error to the next will be small even when re-iterated B times.*

At this stage, we have the tools to address the preceding remark and continue the solution, thus far constructed locally in time, in a global fashion. In fact, working with the presentation (7.66) on an interval $[t_1^n - T_*, t_1^n + R_1]$ for some very large R_1 (whose choice will depend on the scattering properties of $\epsilon_{1, bounded}^{< K_*}$) and essentially repeating the arguments in the proof of Lemma 7.21, picking n large enough, we can extend the solution in the form (7.66) such that $\left\| \tilde{\eta}^{Kn, < K_*} \right\|_{S([t_1^n - T_*, t_1^n + R_1])} \ll \delta_2$. Moreover, we still have the relation

$$\lim_{R, T_*, K, n \rightarrow \infty} \left\| \nabla_{t,x} (\tilde{\eta}_{(1)}^{Kn, < K_*} \Big|_{t_1^n + R_1} - \tilde{\eta}^{Kn, < K_*} \Big|_{t_1^n + R_1}) \right\|_{L_x^2} = 0.$$

Then we construct the solution on the interval $[t_1^n + R_1, t_2^n - T_{1*}]$ for R_1, T_{1*} sufficiently large, by repeating the procedure in Step 2 of Stage 1 before. More precisely, we modify (7.66) to

$$(7.68) \quad \phi^{n, < K_*} = u^n + \sum_k \tilde{P}_k \left((U_{< k}^{(\phi^{n, < K_*})})^\dagger \left(\sum_{b=0}^B \tilde{\epsilon}_{b,k}^{Kn, < K_*} + \tilde{\epsilon}_{B,k}^{Kn, < K_*} \right) \right) + \tilde{\eta}^{Kn, < K_*}$$

on $[t_1^n + R_1, t_2^n - T_{1*}]$, where now $\epsilon_{1, bounded}^{< K_*}$ has been replaced by the expression $\sum_k (U_{< k}^{(\phi^{n, < K_*})})^\dagger \tilde{\epsilon}_{1,k}^{Kn, < K_*}$, where $\tilde{\epsilon}_{1,k}^{Kn, < K_*}$ solves (7.34) with data at time $t = t_1^n + R_1$ given by

$$\tilde{\epsilon}_{1,k}^{Kn, < K_*} [t_1^n + R_1] = \left((U_{< k}^{(\phi^{n, < K_*})})^\dagger \epsilon_{1, bounded}^{< K_*} \right) \Big|_{t_1^n + R_1}, \quad \partial_t (U_{< k}^{(\phi^{n, < K_*})})^\dagger \epsilon_{1, bounded}^{< K_*} \Big|_{t_1^n + R_1}.$$

Then repeating the arguments in the proof of Lemma 7.22, we infer the global bound

$$\left\| \tilde{\eta}^{Kn, < K_*} \right\|_{S([t_1^n + R_1, t_2^n - T_{1*}])} \ll \delta_2.$$

Moreover, importantly, we also obtain that

$$\lim_{R, T_*, R_1, T_{1*}, K, n \rightarrow \infty} \left\| \tilde{\eta}^{Kn, < K_*} - \tilde{\eta}_{(1)}^{Kn, < K_*} \right\|_{S([t_1^n + R_1, t_2^n - T_{1*}])} = 0.$$

Combining this with our earlier considerations, we observe that the error function consisting of $\tilde{\eta}^{Kn, < K_*}$ on $[R, t_1^n - T_*]$, $\tilde{\eta}^{Kn, < K_*}$ on $[t_1^n - T_*, t_1^n + R_1]$ and $\tilde{\eta}^{Kn, < K_*}$ on $[t_1^n + R_1, t_2^n - T_{1*}]$ differs from $\tilde{\eta}_{(1)}^{Kn, < K_*}$ (which we recall solves (7.67) and had its data prescribed at time $t = R$) by an error with respect to $\|\cdot\|_S$, which vanishes as the parameters $R, T_*, R_1, T_{1*}, K, n \rightarrow \infty$.

However, on account of the fact that $\tilde{\epsilon}_B^{Kn, < K_*}$ satisfies uniform S -bounds (its energy being uniformly bounded), the equation (7.67) can be solved for $\tilde{\eta}_{(1)}^{Kn, < K_*}$ via a simple divisibility argument *on any time*

interval on which $\phi^{n, < K_*}$ is defined, and picking $B = B(\|u^n\|_S, \delta_2)$ large enough, we can ensure that

$$\|\tilde{\eta}_{(1)}^{K_n, < K_*}\|_S \ll \delta_2.$$

We can then repeat the preceding procedure and prolong the solution beyond $t_2^n - T_{1*}$, until after B steps we have a global solution. Of course at each of the B many steps, we have to adjust the parameters $R, T_*, R_1, T_{1*}, K, n$ etc. to preserve the required smallness of the error. This proves Theorem 7.17. \square

7.5. Conclusion of the induction on frequency process. In the previous subsection we established that the data $\Pi_{\leq \mu_n^1 + \log(R_n)} \phi^n[0]$ can be globally evolved with uniform S -norm bounds under the assumption that all profiles have energy strictly less than E_{crit} . We may now continue this induction on frequency process and by proceeding as in Subsection 7.3 obtain that the data $\Pi_{\leq \mu_n^2 - \log(R_n)} \phi^n[0]$ can be globally evolved with uniform S -norm bounds. Then we “add in” the second frequency atom $P_{[\mu_n^2 - \log(R_n), \mu_n^2 + \log(R_n)]} \phi^n[0]$ in the sense that by proceeding analogously to Subsection 7.4 we may establish the global evolution of the corresponding geometric data $\Pi_{\leq \mu_n^2 + \log(R_n)} \phi^n[0]$ with uniform S -norm bounds under the assumption that the associated profiles all have energy strictly less than E_{crit} .

We may continue this procedure Λ_0 many times and establish the global evolution with uniform S -norm bounds of the essentially singular sequence of data $\{\phi^n[0]\}_{n \geq 1}$, which would however be a contradiction, unless the sequence $\{\phi^n[0]\}_{n \geq 1}$ is composed of exactly one frequency atom that consists of exactly one profile of asymptotic energy E_{crit} . Thus, in view of Theorem 7.17, we may assume after rescaling so that $\phi^n[0]$ is essentially supported at frequency ~ 1 , that we can either write

$$(7.69) \quad \Pi_{< K_*} \phi^n[0] = \phi_{bounded}^{< K_*}[0] + \eta^{n, < K_*}[0]$$

with $\lim_{K_*, n \rightarrow \infty} \|\eta^{n, < K_*}[0]\|_{\dot{H}_x^1 \times L_x^2} = 0$, or else we have

$$(7.70) \quad \Pi_{< K_*} \phi^n[0] = \sum_{k \in \mathbb{Z}} (U_{< k}^{(\phi^{n, < K_*})})^\dagger S(\cdot - t_n)(\phi_k^{< K_*}[0])[0] + \eta^{n, < K_*}[0]$$

with $\lim_{n \rightarrow \infty} |t_n| = +\infty$, where $S(\cdot)$ now simply denotes the free wave propagator associated with the standard d'Alembertian on \mathbb{R}^{1+2} , and we again have $\lim_{K_*, n \rightarrow \infty} \|\eta^{n, < K_*}[0]\|_{\dot{H}_x^1 \times L_x^2} = 0$.

In any case we infer that the limits

$$\phi_{bounded}[0] = \lim_{K_* \rightarrow \infty} \phi_{bounded}^{< K_*}[0], \quad \phi[0] = \lim_{K_* \rightarrow \infty} \phi^{< K_*}[0]$$

exist and satisfy the “criticality condition”

$$E[\phi_{bounded}] = E[\phi] = E_{crit}.$$

We also observe that we have $\eta^{n, < K_*} \rightarrow 0$ in L_{loc}^∞ as $n \rightarrow \infty$ after passing to subsequences.

In the next two propositions we infer that we can actually extract a minimal blowup solution from either scenario.

Proposition 7.25. *Assume the situation given by (7.69). Then the wave maps evolution $\phi_{bounded}(t, x)$ with lifespan I associated with the energy class data $\phi_{bounded}[0]: \mathbb{R}^2 \rightarrow T\mathbb{S}^m$ satisfies*

$$\sup_{[T_1, T_2] \subset I} \|\phi_{bounded}\|_{S([T_1, T_2])} = +\infty.$$

In particular, $\phi_{bounded}(t, x)$ is a minimal blowup solution.

Proof. We must show that it holds that

$$\lim_{K \rightarrow \infty} \sup_{[T_1, T_2] \subset I} \|\phi_{bounded}^{< K_*}\|_{S([T_1, T_2])} = +\infty.$$

Suppose not. Then all evolutions $\phi_{bounded}^{<K_*}$ exist globally in time and we have that

$$\limsup_{K_* \rightarrow \infty} \|\phi_{bounded}^{<K_*}\|_{S[\mathbb{R}]} < \infty.$$

This uniform bound on the evolutions $\phi_{bounded}^{<K_*}$ implies by Lemma 7.15 that they have uniform scattering behavior. Given $\gamma > 0$, there exists a decomposition

$$\phi_{bounded}^{<K_*} = (\phi_{bounded}^{<K_*})_1 + (\phi_{bounded}^{<K_*})_2$$

and some $R_\gamma > 0$ such that

$$\|(\phi_{bounded}^{<K_*})_1\|_{S[\mathbb{R}]} < \gamma, \quad \|(\phi_{bounded}^{<K_*})_2\|_{L_t^\infty L_x^\infty [(-\infty, -R_\gamma] \cup [R_\gamma, +\infty)]} < \gamma.$$

We can then show, following Stage 1 in the proof of Theorem 7.17, that we obtain uniform (in large enough n and K_*) bounds on $\|\eta^{n, <K_*}\|_{S[\mathbb{R}]}$. This in turn contradicts the fact that for any K_* large enough, we must have $\lim_{n \rightarrow \infty} \|\phi^{n, <K_*}\|_{S[\mathbb{R}]} = +\infty$ on account of Lemma 7.4 and the fact that $\{\phi^n\}_{n \geq 1}$ is an essentially singular sequence which is essentially supported at frequency ~ 1 .

In order to obtain the desired control over $\eta^{n, <K_*}$, we split \mathbb{R}^{1+2} into the regions $[-R_\gamma, R_\gamma]^c \times \mathbb{R}^2$ and $[-R_\gamma, R_\gamma] \times \mathbb{R}^2$. To control $\eta^{n, <K_*}$ on the former region, we use a (much simplified) variant of the argument for Lemma 7.22. To handle the latter region, we argue as in (v) of Step 1 of the proof of Lemma 7.21, replacing $\epsilon_{bounded}^{<K_*}$ there by $\phi_{bounded}^{<K_*}$. \square

To conclude, we have

Proposition 7.26. *Assume the situation given by (7.70). Then there exists an energy class data pair $\tilde{\phi}[0]: \mathbb{R}^2 \rightarrow TS^m$ with $E[\tilde{\phi}] = E_{crit}$ and such that denoting its wave maps evolution with lifespan I by $\tilde{\phi}(t, x)$, we have*

$$\sup_{[T_1, T_2] \subset I} \|\tilde{\phi}\|_{S([T_1, T_2])} = +\infty.$$

In particular, $\tilde{\phi}(t, x)$ is a minimal blowup solution.

Proof. Assume without loss of generality that $t_n \rightarrow +\infty$ as $n \rightarrow \infty$. We reduce to the situation in the preceding Proposition 7.25. To begin with, using the argument of the proof of Lemma 7.22, we see that there is a large $C > 0$ such that the wave maps evolution of the data $\phi^{n, <K_*}[0] \equiv \Pi_{<K_*} \phi^n[0]$ exists on $(-\infty, t^n - C]$ for any large enough n . Then by a simple compactness argument, exploiting the frequency localization up to exponentially decaying tails, we have that

$$\phi^{n, <K_*}[t^n - C] \longrightarrow \tilde{\phi}^{<K_*}[0]$$

as $n \rightarrow \infty$ in $(\dot{H}_{loc}^1 \cap L_{loc}^\infty) \times L_{loc}^2$. We also have that the limit

$$\tilde{\phi}[0] = \lim_{K_* \rightarrow \infty} \tilde{\phi}^{<K_*}[0]$$

exists in $\dot{H}_x^1 \times L_x^2$. We now claim that this is the desired minimal blowup solution. In fact, we can write

$$\phi^{n, <K_*}[t^n - C] = \tilde{\phi}^{<K_*}[0] + \tilde{\eta}^{n, <K_*}[t^n - C],$$

where $\tilde{\eta}^{n, <K_*}(t^n - C) \rightarrow 0$ both in L_{loc}^∞ as well as \dot{H}_{loc}^1 , and also $\partial_t \tilde{\eta}^{n, <K_*}(t^n - C) \rightarrow 0$ in L_{loc}^2 . But then assuming that the evolutions of the $\tilde{\phi}^{<K_*}[0]$ satisfy uniform (in K_*) global-in-time S -norm bounds, we obtain a contradiction exactly as in the proof of the preceding Proposition 7.25. \square

Finally, we observe that the minimal blowup solution constructed in the preceding Proposition 7.25, respectively in Proposition 7.26, has the following crucial compactness property whose proof follows exactly as in [6, 7], see also Corollary 9.36 in [17].

Corollary 7.27. *There exists a non-trivial, radially symmetric, energy class, minimal blowup solution $\phi^\infty: I \times \mathbb{R}^2 \rightarrow \mathbb{S}^m$ to (WM) of energy E_{crit} and with maximal interval of existence I . Moreover, there exists a continuous function $\lambda: I \rightarrow (0, \infty)$ so that the family of functions*

$$\{(\phi^\infty(t, \lambda(t)^{-1} \cdot), \lambda(t)^{-1} \partial_t \phi^\infty(t, \lambda(t)^{-1} \cdot)): t \in I\}$$

is pre-compact in $\dot{H}_x^1 \times L_x^2$.

8. RIGIDITY ARGUMENT

In this last section we rule out the existence of a minimal blowup solution ϕ^∞ as in Corollary 7.27. To this end we closely mimic the rigidity argument of Kenig-Merle [7] as implemented in [17] for energy critical wave maps with \mathbb{H}^2 target. We emphasize that the momentum vanishing property of the minimal blowup solution is here an immediate consequence of the hypothesis of radial symmetry.

8.1. Preliminary properties of minimal blowup solutions. We begin by stating some standard properties of the minimal blowup solution ϕ^∞ . As in Corollary 7.27, I denotes the lifespan of ϕ^∞ and we set $I^+ := I \cap [0, \infty)$. The finite speed of propagation for wave maps gives the following lemma.

Lemma 8.1. *For given $\varepsilon > 0$, let $M > 0$ be such that*

$$(8.1) \quad \int_{|x| \geq M} \frac{1}{2} \sum_{\alpha=0}^2 |\partial_\alpha \phi^\infty(0, x)|^2 dx < \varepsilon.$$

Then

$$(8.2) \quad \int_{|x| \geq M+t} \frac{1}{2} \sum_{\alpha=0}^2 |\partial_\alpha \phi^\infty(t, x)|^2 dx < \varepsilon.$$

for all $t \in I^+$.

Next, we recall the following conservation laws for wave maps.

Proposition 8.2. *Let ϕ be an energy class solution to (WM). Then the following conservation laws hold.*

- *Energy conservation*

$$(8.3) \quad \frac{d}{dt} \int_{\mathbb{R}^2} \frac{1}{2} \sum_{\alpha=0}^2 |\partial_\alpha \phi|^2 dx = 0$$

- *Momentum conservation*

$$(8.4) \quad \frac{d}{dt} \int_{\mathbb{R}^2} \partial_t \phi^\dagger \partial_i \phi dx = 0 \quad i = 1, 2$$

- *Weighted energy*

$$(8.5) \quad \frac{d}{dt} \int_{\mathbb{R}^2} \sum_{\alpha=0}^2 x_i \varphi(x/R) \frac{1}{2} |\partial_\alpha \phi(t, x)|^2 dx = - \int_{\mathbb{R}^2} \partial_t \phi^\dagger \partial_i \phi dx + O(r(R)) \quad i = 1, 2$$

- *Weighted momentum monotonicity*

$$(8.6) \quad \frac{d}{dt} \int_{\mathbb{R}^2} \sum_{i=1}^2 x_i \varphi(x/R) \partial_i \phi^\dagger \partial_i \phi dx = - \int_{\mathbb{R}^2} |\partial_t \phi|^2 dx + O(r(R))$$

where φ is a fixed bump function which is equal to one on $\{|x| \leq 1\}$ and

$$r(R) := \int_{|x| \geq R} \sum_{\alpha=0}^2 |\partial_\alpha \phi^\infty(0, x)|^2 dx.$$

Proof. For classical solutions to (WM) these identities follow from Stokes' theorem together with the fact that the stress-energy tensor is divergence free. By Definition 7.6 one can then pass to the limit and obtain the result for energy class solutions. \square

By virtue of Corollary 7.27, the energy class, minimal blowup solution ϕ^∞ satisfies some important properties, which we present in the next lemmas. In the case that I^+ is finite, we have the following lower bound for the continuous function $\lambda(t)$ of Corollary 7.27.

Lemma 8.3. *Assume that I^+ is finite. After suitable rescaling, we may assume that $I^+ = [0, 1)$. Let $\lambda : I^+ \rightarrow (0, \infty)$ be as in Corollary 7.27 and let us denote the pre-compact set in $\dot{H}_x^1 \times L_x^2$ therein by K . Then, there exists a constant $C_0(K) > 0$ such that*

$$(8.7) \quad 0 < \frac{C_0(K)}{1-t} \leq \lambda(t)$$

for all $0 \leq t < 1$.

Proof. The proof follows along the lines of Lemma 10.4 in [17] \square

In addition, under the same assumptions of Lemma 8.3 we have the following support property of ϕ^∞ .

Lemma 8.4. *Let ϕ^∞ as in Corollary 7.27 with finite lifespan. Then there exists $x_0 \in \mathbb{R}^2$ such that*

$$\text{supp}(\phi^\infty(t, \cdot)) \subset B(x_0, 1-t)$$

for all $0 \leq t < 1$.

Proof. The proof follows as in Lemma 4.8 of [7] by use of Lemmas 8.1 and 8.3. \square

We finally state the vanishing momentum property which we get here for free by the very definition of radially symmetric wave maps.

Proposition 8.5. *Let ϕ^∞ as in Corollary 7.27. Then*

$$(8.8) \quad \int_{\mathbb{R}^2} \partial_t \phi^{\infty \dagger} \partial_i \phi^\infty dx = 0 \quad i = 1, 2$$

for all times in I^+ .

8.2. Rigidity I: Infinite time interval and reduction to the self-similar case for finite time intervals.
The final goal of the rigidity argument consists in proving the following theorem.

Proposition 8.6. *Let ϕ^∞ as in Corollary 7.27 with lifespan $I = (-T_0, T_1)$. Then, one cannot have T_1 or T_0 finite. Moreover, if $\lambda(t) \geq \lambda_0 > 0$ for all $t \in \mathbb{R}$, one necessarily has $\partial_\alpha \phi^\infty = 0$ for $\alpha = 0, 1, 2$.*

In this subsection, we will provide the proof of Proposition 8.6 for the infinite time case and then reduce to the self-similar situation for the finite time case. The proof of Proposition 8.6 in the finite time case will be concluded in the next subsection.

Proof. [Proposition 8.6: infinite time case] We will first treat the case $T_1 = \infty$. Assuming that $\partial_\alpha \phi^\infty$ do not all vanish, we extract a non-trivial finite energy radially symmetric harmonic map into the sphere, leading to a contradiction. Following the proof of Lemma 10.9 in [17], which crucially hinges on the geometry of the target, we show that there exists $\alpha \in \mathbb{R}$ such that

$$(8.9) \quad \int_I \int_{\mathbb{R}^2} |\partial_t \phi^\infty(t, x)|^2 dx dt \geq \alpha > 0$$

for all intervals I of unit length. Suppose not, then there exists a sequence of intervals $I_n := [t_n, t_n + 1]$ with the property that $t_n \rightarrow \infty$ and

$$(8.10) \quad \int_{I_n} \int_{\mathbb{R}^2} |\partial_t \phi^\infty(t, x)|^2 dx dt \leq \frac{1}{n}.$$

Then there exists a sequence of times $s_n \in I_n$ such that $\|\partial_t \phi^\infty(s_n, \cdot)\|_2 \rightarrow 0$ as $n \rightarrow \infty$. By the compactness property of Corollary 7.27, the family of functions

$$\left\{(\phi^\infty(s_n, \lambda(s_n)^{-1} \cdot), \lambda(s_n)^{-1} \partial_t \phi^\infty(s_n, \lambda(s_n)^{-1} \cdot))\right\}_{n=0}^\infty$$

is pre-compact in $\dot{H}_x^1 \times L_x^2$. Up to a subsequence, there exist limit functions Φ^* such that

$$\lambda(s_n)^{-1} \partial_\alpha \phi^\infty(s_n, \lambda(s_n)^{-1} \cdot) \rightarrow \partial_\alpha \Phi^*(\cdot)$$

strongly in L_x^2 for $\alpha = 0, 1, 2$. Pre-compactness and standard perturbative arguments ensure that there exists some non-empty interval I^* around zero such that

$$\lambda(s_n)^{-1} \partial_\alpha \phi^\infty(s_n + t \lambda(s_n)^{-1} \cdot, \lambda(s_n)^{-1} \cdot) \rightarrow \partial_\alpha \Phi^*(t, \cdot)$$

in $L_{loc}^\infty(I^*; L^2(\mathbb{R}^2))$ with Φ^* being a radially symmetric weak solution of (WM) on $I^* \times \mathbb{R}^2$. We now distinguish two cases:

- there exists a sequence of times $s_n \in I_n$ such that $\{\lambda(s_n)\}$ is bounded: since $\lambda(t) \geq \lambda_0 > 0$, there exists a non-empty time interval $\tilde{I} \subset I^*$ such that $s_n + \lambda(s_n)^{-1} \tilde{I} \subset I_n$ for each n ; then (8.10) implies that $\int_{\tilde{I}} \int_{\mathbb{R}^2} |\partial_t \Phi^*(t, x)|^2 dx dt = 0$ whence $\partial_t \Phi^*(t, \cdot) = 0$ for all $t \in \tilde{I}$.
- $\{\lambda(s_n)\}$ is unbounded for every sequence $\{s_n\}$ with $s_n \in I_n$: in this case we use a crucial Vitali covering argument as in [17] to arrive at the same conclusion; we write for each n

$$I_n = \bigcup_{s \in I_n} [s - \lambda(s)^{-1}, s + \lambda(s)^{-1}] \cap I_n$$

and by the Vitali covering lemma, we may pick a disjoint subcollection of intervals $\{I_s\}_{s \in J_n}$ with $I_s := [s - \lambda(s)^{-1}, s + \lambda(s)^{-1}] \cap I_n$ for some subset $J_n \subset I_n$ with the property that

$$\sum_{s \in J_n} 5 |I_s| \geq 1;$$

it follows that, by definition of I_n , we may pick a sequence of times $s_n \in I_n$ with the property that $\int_{I_{s_n}} \|\partial_t \phi^\infty(t, \cdot)\|_{L_x^2}^2 dt = o(\lambda^{-1}(s_n))$ and in particular

$$\int_{-1}^1 \|(\chi_{I_n} \partial_t \phi^\infty)(s_n + t \lambda(s_n)^{-1} \cdot, \cdot)\|_{L_x^2}^2 dt = o(1),$$

but then, passing to a subsequence, we can extract a limiting function Φ^* from

$$\lambda(s_n)^{-1} \partial_t \phi^\infty(s_n + t \lambda(s_n)^{-1} \cdot, \lambda(s_n)^{-1} \cdot)$$

with the property that $\partial_t \Phi^*(t, \cdot) = 0$ for all t in the lifespan of Φ^* .

Hence, in both cases we have deduced the existence of a radially symmetric weak wave map $\Phi^* : \mathbb{R}^{2+1} \rightarrow \mathbb{S}^2$ with the following properties

$$\partial_t \Phi^* = 0, \quad t \in I^*$$

and

$$\sum_{\alpha=1}^2 \|\partial_\alpha \Phi^*\|_{L_x^2}^2 \neq 0.$$

This leads to a contradiction since there do not exist non-vanishing finite energy radial harmonic maps into the sphere, see for instance [31]. We can then conclude that (8.9) holds. The remainder of the argument hinges upon the weighted momentum monotonicity (8.6) combined with the following property: for any given $\varepsilon > 0$, by pre-compactness of the orbit of ϕ^∞ in $\dot{H}_x^1 \times L_x^2$ there exists $R(\varepsilon) > 0$ such that for all $t \in I$,

$$\int_{\{|x| \geq R(\varepsilon)\}} |\lambda(t)^{-1} \nabla_{t,x} \phi^\infty(t, \lambda(t)^{-1} x)|^2 dx \leq \varepsilon.$$

Changing variables and using the key assumption $\lambda(t) \geq \lambda_0 > 0$, we can conclude that for all $t \in I$,

$$\int_{\{|x| \geq \frac{R(\varepsilon)}{\lambda_0}\}} |\nabla_{t,x} \phi^\infty(t, x)|^2 dx \leq \varepsilon.$$

Using the notation $\tilde{R}(\varepsilon) := \frac{R(\varepsilon)}{\lambda_0}$, then by the weighted momentum monotonicity identity (8.6) integrated over a sufficiently large time interval $[0, T]$, $T \gg 1$, we get for arbitrary $\varepsilon > 0$ and corresponding $\tilde{R} \equiv \tilde{R}(\varepsilon)$ that

$$(8.11) \quad \int_{\mathbb{R}^2} x^j \varphi(x/\tilde{R}) (\partial_t \phi^\infty)^\dagger \partial_j \phi^\infty dx \Big|_0^T = - \int_0^T \int_{\mathbb{R}^2} |\partial_t \phi^\infty(t, x)|^2 dx dt + \int_0^T O(r(\tilde{R})) dt,$$

where

$$r(\tilde{R}) = \int_{|x| \geq \tilde{R}} |\nabla_{t,x} \phi^\infty(t, x)|^2 dx \leq \varepsilon$$

by our choice of \tilde{R} . But then using (8.9), the right-hand side of (8.11) can be bounded from above by

$$- \int_0^T \int_{\mathbb{R}^2} |\partial_t \phi^\infty(t, x)|^2 dx dt + \int_0^T O(r(\tilde{R})) dt \leq -(T-1)\alpha + CT\varepsilon \longrightarrow -\infty$$

as $T \rightarrow \infty$, if we just choose $\varepsilon > 0$ sufficiently small so that $C\varepsilon \ll \alpha$. At this point $\varepsilon > 0$ and correspondingly \tilde{R} are fixed. But then we see that the left-hand side of (8.11) is bounded by $\tilde{R}E_{crit}$ and so we get a contradiction by taking $T \gg 1$ sufficiently large. \square

Having concluded the infinite time case, we turn now to the proof of Proposition 8.6 in the case $T_1 < \infty$. Similarly to [7], [17] and [16], the first step consists in reducing to a self-similar blow-up scenario. As before, we can assume without loss of generality that $T_1 = 1$. We recall from Lemma 8.3 that

$$(8.12) \quad \lambda(t) \geq \frac{C_0(K)}{1-t}, \quad 0 < t < 1,$$

and from Lemma 8.4 that

$$(8.13) \quad \text{supp}(\phi^\infty(t, \cdot)) \subset B(0, 1-t) \quad \text{and} \quad \text{supp}(\partial_t \phi^\infty(t, \cdot)) \subset B(0, 1-t).$$

Then we may prove the following upper bound for $\lambda(t)$.

Lemma 8.7. *Let ϕ^∞ as above with $T_1 = 1$. Then there exists $C_1(K) > 0$ such that*

$$(8.14) \quad \lambda(t) \leq \frac{C_1(K)}{1-t}$$

for all $0 \leq t < 1$.

Proof. We adopt the same strategy as in the proof of Lemma 10.11 in [17] (see also Lemma 8.12 in [16]). Suppose that (8.14) fails. Define, for $0 \leq t < 1$ the following functional

$$z(t) := \sum_{j=1}^2 \int_{\mathbb{R}^2} x_j (\partial_j \phi^\infty(t, x))^\dagger \partial_t \phi^\infty(t, x) dx.$$

The weighted momentum monotonicity (8.6) allows to compute that

$$z'(t) = - \int_{\mathbb{R}^2} |\partial_t \phi^\infty(t, x)|^2 dx.$$

Next by (8.13), we see that $z(t) \rightarrow 0$ as $t \rightarrow 1$, hence we can write

$$z(t) = \int_t^1 \int_{\mathbb{R}^2} |\partial_t \phi^\infty(s, x)|^2 dx ds.$$

At this point, we need to distinguish two cases: either there exists $\alpha > 0$ such that

$$\int_t^1 \int_{\mathbb{R}^2} |\partial_t \phi^\infty(s, x)|^2 dx ds \geq \alpha(1-t), \quad 0 \leq t < 1$$

or else, there exists a sequence $\{t_n\}_n \subset [0, 1)$ with $t_n \rightarrow 1$ such that, denoting $J_n = (t_n, 1)$, it holds

$$|J_n| \int_{J_n} \int_{\mathbb{R}^2} |\partial_t \phi^\infty(s, x)|^2 dx ds \rightarrow \infty \quad \text{as } n \rightarrow \infty.$$

In the first case, we get a contradiction by proceeding as in the proof of Lemma 5.6 of [7] and using therein the vanishing momentum property of Proposition 8.5. In the second case, we argue similarly as in the proof of Proposition 8.6 for the infinite time case: by a Vitali argument and by pre-compactness one can conclude the existence of a non-vanishing finite energy radially symmetric harmonic map into the sphere which gives a contradiction (cf. [31]). \square

Combining the previous lemmas we can reduce to the self-similar scenario. Indeed, we have the following.

Corollary 8.8. *Let ϕ^∞ be as above with $T_1 = 1$, then the set*

$$\{(1-t)\partial_\alpha \phi^\infty(t, (1-t)\cdot) : t \in I, \alpha = 0, 1, 2\}$$

is pre-compact in $L_x^2(\mathbb{R}^2)$.

Proof. For the proof we refer to the proof of Proposition 5.7 in [7]. \square

8.3. Rigidity II: The self-similar case. In this section we deal with the finite time case and we carry out the last step of the Kenig-Merle rigidity argument by excluding the possibility of self-similar blow-up. In particular, we rule out the existence of a minimal blow-up solution ϕ^∞ as in Corollary 8.8. For the sake of simplicity we drop the superscript ∞ and denote the minimal blowup solution from Corollary 8.8 just by ϕ . As in [20], [7], [17], we introduce the self-similar variables

$$y = \frac{x}{1-t}, \quad s = -\log(1-t), \quad 0 \leq t < 1$$

and

$$\widetilde{\phi}(s, y, 0) := \phi(t, x) = \phi(1 - e^{-s}, e^{-s}y), \quad 0 \leq s < \infty.$$

By construction, $\nabla_{s,y} \widetilde{\phi}(s, \cdot, 0)$ is supported in $\{y \in \mathbb{R}^2 : |y| \leq 1\}$. Next, for small $\delta > 0$, we also define

$$y = \frac{x}{1+\delta-t}, \quad s = -\log(1+\delta-t), \quad 0 \leq t < 1$$

and set

$$(8.15) \quad \widetilde{\phi}(s, y, \delta) := \phi(t, x) = \phi(1 + \delta - e^{-s}, e^{-s}y), \quad 0 \leq s < \infty.$$

Note that $\widetilde{\phi}(s, y, \delta)$ is defined for $-\log(1+\delta) \leq s < -\log \delta$. By standard computations, we can write the wave maps equation in self-similar variables as follows

$$(8.16) \quad \partial_s^2 \widetilde{\phi} = \frac{1}{\rho} \div (\rho \nabla_y \widetilde{\phi} - \rho(y \cdot \nabla_y \widetilde{\phi})y) - 2y \cdot \nabla_y \partial_s \widetilde{\phi} - \partial_s \widetilde{\phi} + \widetilde{\phi} \left(|\partial_s \widetilde{\phi} + y \cdot \nabla_y \widetilde{\phi}|^2 - |\nabla_y \widetilde{\phi}|^2 \right)$$

where $\rho = (1 - |y|^2)^{-1/2}$. The following lemma collects some basic properties of $\widetilde{\phi}$ inherited from ϕ .

Lemma 8.9. *For fixed $\delta > 0$, we have for all $0 \leq s < -\log \delta$ that*

- $\text{supp}(\partial_\alpha \widetilde{\phi}(s, \cdot, \delta)) \subset \{y \in \mathbb{R}^2 : |y| \leq 1 - \delta\} \quad \alpha = 0, 1, 2$
- $\int_{\mathbb{R}^2} \left(|\nabla_y \widetilde{\phi}(s, y, \delta)|^2 + |\partial_s \widetilde{\phi}(s, y, \delta)|^2 \right) dy \leq C$
- $\sum_{\alpha=0}^2 \int_{\mathbb{R}^2} |\partial_\alpha \widetilde{\phi}(s, y, \delta)|^2 \log \left(\frac{1}{1-|y|^2} \right) dy \leq C \log \frac{1}{\delta}$
- $\sum_{\alpha=0}^2 \int_{\mathbb{R}^2} |\partial_\alpha \widetilde{\phi}(s, y, \delta)|^2 (1 - |y|^2)^{-\frac{1}{2}} dy \leq C \delta^{-1/2}.$

Proof. The properties are obtained by direct calculation. See also [7] and [16]. \square

As in [17], we introduce the following Lyapunov functional

$$\widetilde{E}(\widetilde{\phi})(s) := \frac{1}{2} \int_{B_1} \left[|\partial_s \widetilde{\phi}|^2 + |\nabla_y \widetilde{\phi}|^2 - |y \cdot \nabla_y \widetilde{\phi}|^2 \right] (1 - |y|^2)^{-\frac{1}{2}} dy,$$

which satisfies a suitable monotonicity property stated in the next proposition.

Proposition 8.10. *For $0 \leq s_1 < s_2 < \log(1/\delta)$, the following identities holds*

- (1) $\widetilde{E}(\widetilde{\phi})(s_2) - \widetilde{E}(\widetilde{\phi})(s_1) = \int_{s_1}^{s_2} \int_{B_1} \frac{|\partial_s \widetilde{\phi}|^2}{(1 - |y|^2)^{3/2}} dy ds,$
- (2) $\lim_{s \rightarrow \log(1/\delta)} \widetilde{E}(\widetilde{\phi})(s) \leq E_{crit}.$

Proof. For (1) see the proof of Lemma 2.1 in [20], while (2) can be proven as in [7]. \square

By direct application of Proposition 8.10 we obtain the following corollary.

Lemma 8.11. *For all $\delta > 0$, there exists $\bar{s}_\delta \in \left(\frac{|\log \delta|}{2}, |\log \delta|\right)$ such that*

$$\int_{\bar{s}_\delta}^{\bar{s}_\delta + |\log \delta|^{\frac{1}{2}}} \int_{B_1} \frac{|\partial_s \widetilde{\phi}|^2}{(1 - |y|^2)^{3/2}} dy ds \leq \frac{E_{crit}}{|\log \delta|^{\frac{1}{2}}}.$$

By using once more pre-compactness together with the self-similar behavior, we aim to extract a limit ϕ^* as $\delta \rightarrow 0$ and to show that ϕ^* is a stationary solution to (8.16) and finally obtain a contradiction as in previous sections. In order to achieve this, let $\bar{t}_\delta = 1 + \delta - e^{-\bar{s}_\delta}$ for \bar{s}_δ as in Lemma 8.11. By Corollary 8.8 we may select $\delta_j \rightarrow 0$ as $j \rightarrow \infty$ such that, for each $\alpha = 0, 1, 2$

$$(1 - \bar{t}_{\delta_j}) \partial_\alpha \phi^\infty(\bar{t}_{\delta_j}, (1 - \bar{t}_{\delta_j})x) \rightarrow \partial_\alpha \Phi^*(x)$$

strongly in L^2 as $\delta_j \rightarrow 0$. In fact, we may also achieve that

$$(8.17) \quad (1 + \delta_j - \bar{t}_{\delta_j}) \partial_\alpha \phi^\infty(\bar{t}_{\delta_j}, (1 + \delta_j - \bar{t}_{\delta_j})x) \rightarrow \partial_\alpha \Phi^*(x)$$

strongly in L^2 . Next, we consider the evolution, in the sense of Definition 7.6, of the energy class data given by the left hand side of (8.17). We denote these evolutions by ϕ^{j*} and we remark that, due to pre-compactness and standard perturbative arguments, these evolutions exist on some joined fixed life-span $[0, T^*]$, where we may assume $0 < T^* < 1$. Furthermore, on $[0, T^*]$ we have

$$\partial_\alpha \phi^{j*}(t, x) = (1 + \delta_j - \bar{t}_{\delta_j}) \partial_\alpha \phi^\infty(\bar{t}_{\delta_j} + (1 + \delta_j - \bar{t}_{\delta_j})t, (1 + \delta_j - \bar{t}_{\delta_j})x)$$

and for each $\alpha = 0, 1, 2$

$$\partial_\alpha \phi^{j*}(t, \cdot) \rightarrow \partial_\alpha \Phi^*(t, \cdot)$$

strongly in L^2 as $j \rightarrow \infty$ and uniformly for all $0 \leq t \leq T^*$, where Φ^* is a weak wave map on $[0, T^*] \times \mathbb{R}^2$. We also remark that, due to the previous identities, it holds

$$\text{supp}(\phi^{j*}(t, \cdot)) \subset \left\{ x \in \mathbb{R}^2 : |x| \leq \frac{1 - \bar{t}_{\delta_j}}{1 + \delta_j - \bar{t}_{\delta_j}} - t < 1 - t \right\}$$

for $0 \leq t \leq T^*$. Next, we switch to self-similar variables

$$y = \frac{x}{1 - t}, \quad s = -\log(1 - t), \quad 0 \leq t \leq T^*,$$

and define

$$\widetilde{\phi}^{j*}(s, y) := \phi^{j*}(1 - e^{-s}, e^{-s}y)$$

and analogously for $\widetilde{\Phi}^*$. Then, as in [17], we infer that

$$\partial_\alpha \widetilde{\phi}^{j*}(s, \cdot) \rightarrow \partial_\alpha \widetilde{\Phi}^*(s, \cdot)$$

strongly in $L_y^2(\mathbb{R}^2)$ as $j \rightarrow \infty$ and uniformly for all $0 \leq s \leq -\log(1 - T^*/2) =: \widetilde{T}$ and for $\alpha = 0, 1, 2$. Moreover, with $\widetilde{\phi}$ as in (8.15), we have

$$\widetilde{\phi}^{j*}(s, y) = \widetilde{\phi}(\bar{s}_{\delta_j} + s, y, \delta_j)$$

and hence for each $\alpha = 0, 1, 2$

$$(8.18) \quad \partial_\alpha \widetilde{\phi}(\bar{s}_{\delta_j} + s, y, \delta_j) \rightarrow \partial_\alpha \widetilde{\Phi}^*(s, \cdot)$$

strongly in L_y^2 and uniformly in $0 \leq s \leq \widetilde{T}$ as $j \rightarrow \infty$. Therefore, Φ^* is a solution of (8.16) and

$$\text{supp}(\partial_\alpha \widetilde{\Phi}^*(\cdot, s)) \subset \{y \in \mathbb{R}^2 : |y| \leq 1\}$$

and moreover

$$\text{trace}(\widetilde{\Phi}^*(\cdot, s)) = \text{const}$$

where “trace” is the L^2 -trace.

Lemma 8.12. *Let $\widetilde{\Phi}^*$ be as above. Then we have*

$$\widetilde{\Phi}^*(s, y) = \widetilde{\Phi}^*(y) \quad \text{and} \quad \widetilde{\Phi}^* \neq \text{const}.$$

Proof. The proof works exactly as in [17]: by (8.18) and Lemma 8.11 one shows that $\widetilde{\Phi}^*(s, y) = \widetilde{\Phi}^*(y)$ while the fact that $\widetilde{\Phi}^* \neq \text{const}$ follows as in [7]. \square

By this procedure we have obtained a stationary, nonconstant, distributional solution to (8.16) with finite energy (relative to the y variable). As in [17], the following proposition leads to a contradiction.

Proposition 8.13. *Let Φ be a distributional solution to (8.16) of finite energy*

$$\int_{B_1} |\nabla \Phi(y)|^2 dy < \infty.$$

Then $\Phi^ = \text{const}$.*

Proof. For the proof we refer the reader to the proof of Proposition 10.17 in [17] where the argument of Section 7.5.1 in [26] is implemented and the conclusion is reached by Lemaire’s uniqueness theorem [19]. \square

Thanks to Proposition 8.13 we can conclude the proof of Proposition 8.6 in the finite time case.

Proof. [Proposition 8.6: finite time case] In the finite time case, thanks to Corollary 8.8 we have been able to construct a nonconstant self-similar solution Φ^* of (8.16) of finite energy, but this is in contradiction with Proposition 8.13 and this completes the proof of Proposition 8.6. \square

In order to close the rigidity argument, we notice that we need to justify the additional condition $\lambda(t) \geq \lambda_0 > 0$ for all $t \in \mathbb{R}$ assumed in Proposition 8.6, but this follows as in Lemma 10.18 of [17]. This concludes the rigidity argument and hence finishes the proof of Theorem 1.1.

REFERENCES

1. H. Bahouri and P. Gérard, *High frequency approximation of solutions to critical nonlinear wave equations*, Amer. J. Math. **121** (1999), no. 1, 131–175.
2. J. Bourgain, *Estimates for cone multipliers*, Geometric aspects of functional analysis (Israel, 1992–1994), Oper. Theory Adv. Appl., vol. 77, Birkhäuser, Basel, 1995, pp. 41–60.
3. D. Christodoulou and S. Tahvildar-Zadeh, *On the asymptotic behavior of spherically symmetric wave maps*, Duke Math. J. **71** (1993), no. 1, 31–69.
4. ———, *On the regularity of spherically symmetric wave maps*, Comm. Pure Appl. Math. **46** (1993), no. 7, 1041–1091.
5. D. Fang and C. Wang, *Some remarks on Strichartz estimates for homogeneous wave equation*, Nonlinear Anal. **65** (2006), no. 3, 697–706.

6. C. Kenig and F. Merle, *Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case*, Invent. Math. **166** (2006), no. 3, 645–675.
7. ———, *Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation*, Acta Math. **201** (2008), no. 2, 147–212.
8. S. Klainerman and M. Machedon, *Space-time estimates for null forms and the local existence theorem*, Comm. Pure Appl. Math. **46** (1993), no. 9, 1221–1268.
9. ———, *Smoothing estimates for null forms and applications*, Duke Math. J. **81** (1995), no. 1, 99–133 (1996).
10. ———, *On the regularity properties of a model problem related to wave maps*, Duke Math. J. **87** (1997), no. 3, 553–589.
11. S. Klainerman and I. Rodnianski, *On the global regularity of wave maps in the critical Sobolev norm*, Internat. Math. Res. Notices (2001), no. 13, 655–677.
12. S. Klainerman and S. Selberg, *Remark on the optimal regularity for equations of wave maps type*, Comm. Partial Differential Equations **22** (1997), no. 5-6, 901–918.
13. ———, *Bilinear estimates and applications to nonlinear wave equations*, Commun. Contemp. Math. **4** (2002), no. 2, 223–295.
14. J. Krieger, *Global regularity of wave maps from \mathbb{R}^{3+1} to surfaces*, Comm. Math. Phys. **238** (2003), no. 1-2, 333–366.
15. ———, *Global regularity of wave maps from \mathbb{R}^{2+1} to \mathbb{H}^2 . Small energy*, Comm. Math. Phys. **250** (2004), no. 3, 507–580.
16. J. Krieger and J. Lührmann, *Concentration Compactness for the Critical Maxwell-Klein-Gordon Equation*, Annals of PDE **1** (2015), no. 1, 1–208.
17. J. Krieger and W. Schlag, *Concentration Compactness for Critical Wave Maps*, EMS Monographs in Mathematics, European Mathematical Society (EMS), Zürich, 2012.
18. J. Krieger, W. Schlag, and D. Tataru, *Renormalization and blow up for charge one equivariant critical wave maps*, Invent. Math. **171** (2008), no. 3, 543–615.
19. L. Lemaire, *Applications harmoniques de surfaces riemanniennes*, J. Differential Geom. **13** (1978), no. 1, 51–78.
20. F. Merle and H. Zaag, *Determination of the blow-up rate for a critical semilinear wave equation*, Math. Ann. **331** (2005), no. 2, 395–416.
21. G. Métivier and S. Schochet, *Trilinear resonant interactions of semilinear hyperbolic waves*, Duke Math. J. **95** (1998), no. 2, 241–304.
22. J. Nahas, *Scattering of wave maps from \mathbb{R}^{2+1} to general targets*, Calc. Var. Partial Differential Equations **46** (2013), no. 1-2, 427–437.
23. A. Nahmod, A. Stefanov, and K. Uhlenbeck, *On the well-posedness of the wave map problem in high dimensions*, Comm. Anal. Geom. **11** (2003), no. 1, 49–83.
24. P. Raphaël and I. Rodnianski, *Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang-Mills problems*, Publ. Math. Inst. Hautes Études Sci. **115** (2012), 1–122.
25. I. Rodnianski and J. Sterbenz, *On the formation of singularities in the critical $O(3)$ σ -model*, Ann. of Math. (2) **172** (2010), no. 1, 187–242.
26. J. Shatah and M. Struwe, *Geometric wave equations*, Courant Lecture Notes in Mathematics, vol. 2, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1998.
27. ———, *The Cauchy problem for wave maps*, Int. Math. Res. Not. (2002), no. 11, 555–571.
28. J. Sterbenz, *Angular regularity and Strichartz estimates for the wave equation*, Int. Math. Res. Not. (2005), no. 4, 187–231, With an appendix by Igor Rodnianski.
29. J. Sterbenz and D. Tataru, *Energy dispersed large data wave maps in $2 + 1$ dimensions*, Comm. Math. Phys. **298** (2010), no. 1, 139–230.
30. ———, *Regularity of wave-maps in dimension $2 + 1$* , Comm. Math. Phys. **298** (2010), no. 1, 231–264.
31. M. Struwe, *Radially symmetric wave maps from $(1 + 2)$ -dimensional Minkowski space to the sphere*, Math. Z. **242** (2002), no. 3, 407–414.
32. ———, *Equivariant wave maps in two space dimensions*, Comm. Pure Appl. Math. **56** (2003), no. 7, 815–823.
33. ———, *Radially symmetric wave maps from $(1 + 2)$ -dimensional Minkowski space to general targets*, Calc. Var. Partial Differential Equations **16** (2003), no. 4, 431–437.
34. T. Tao, *Global regularity of wave maps. III – VII*, arXiv preprints.
35. ———, *Endpoint bilinear restriction theorems for the cone, and some sharp null form estimates*, Math. Z. **238** (2001), no. 2, 215–268.
36. ———, *Global regularity of wave maps. I. Small critical Sobolev norm in high dimension*, Internat. Math. Res. Notices (2001), no. 6, 299–328.
37. ———, *Global regularity of wave maps. II. Small energy in two dimensions*, Comm. Math. Phys. **224** (2001), no. 2, 443–544.
38. D. Tataru, *On global existence and scattering for the wave maps equation*, Amer. J. Math. **123** (2001), no. 1, 37–77.
39. ———, *Rough solutions for the wave maps equation*, Amer. J. Math. **127** (2005), no. 2, 293–377.
40. G. N. Watson, *A treatise on the theory of Bessel functions*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1995, Reprint of the second (1944) edition.
41. T. Wolff, *A sharp bilinear cone restriction estimate*, Ann. of Math. (2) **153** (2001), no. 3, 661–698.

BÂTIMENT DES MATHÉMATIQUES, EPFL, STATION 8, 1015 LAUSANNE, SWITZERLAND
E-mail address: `elisabetta.chiodaroli@epfl.ch`

BÂTIMENT DES MATHÉMATIQUES, EPFL, STATION 8, 1015 LAUSANNE, SWITZERLAND
E-mail address: `joachim.krieger@epfl.ch`

DEPARTMENT OF MATHEMATICS, JOHNS HOPKINS UNIVERSITY, 3400 N. CHARLES STREET, BALTIMORE, MD 21218, USA
E-mail address: `luehrmann@math.jhu.edu`